

1231698

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

September 29, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM
THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK
OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT
APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A
FILING DATE.

APPLICATION NUMBER: 60/575,429

FILING DATE: May 28, 2004

RELATED PCT APPLICATION NUMBER: PCT/US04/16614

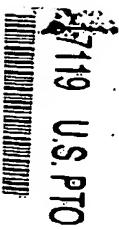
Certified by



Jon W. Dudas

Acting Under Secretary of Commerce
for Intellectual Property
and Acting Director of the U.S.
Patent and Trademark Office





PTO/SB/16 (08-03)

Approved for use through 07/31/2006, OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. ER635153476US

GPO
500-57429
22676

032804

INVENTOR(S)			
Given Name (first and middle [if any]) Soumitra	Family Name or Surname Roy	Residence (City and either State or Foreign Country) Wayne, Pennsylvania	
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto			
TITLE OF THE INVENTION (500 characters max) Methods of Generating Chimeric Adenoviruses and Uses for Such Chimeric Adenoviruses			
Direct all correspondence to: CORRESPONDENCE ADDRESS			
<input checked="" type="checkbox"/> Customer Number: <div style="border: 1px solid black; padding: 2px; width: 150px; margin-left: 10px;">00270</div>			
OR			
<input type="checkbox"/> Firm or Individual Name Address	Howson and Howson Cathy A. Kodroff		
Address	Spring House Corporate Center, 321 Norristown Road, Suite 200		
City	Spring House	State	PA
Country	USA	Telephone	215-540-9200
		Zip	19477
		Fax	215-540-5818
ENCLOSED APPLICATION PARTS (check all that apply)			
<input checked="" type="checkbox"/> Specification Number of Pages <u>64</u> <input checked="" type="checkbox"/> Drawing(s) Number of Sheets <u>2</u> <input checked="" type="checkbox"/> Application Data Sheet. See 37 CFR 1.76		<input type="checkbox"/> CD(s), Number _____ <input checked="" type="checkbox"/> Other (specify) <u>159 pp., Sequence Listing</u>	
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT			
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. <input checked="" type="checkbox"/> A check or money order is enclosed to cover the filing fees. <input checked="" type="checkbox"/> The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: <u>08-3040</u> <input type="checkbox"/> Payment by credit card. Form PTO-2038 is attached.		FILING FEE Amount (\$) <div style="border: 1px solid black; padding: 2px; width: 100px; text-align: center;">160.00</div>	
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.			
<input checked="" type="checkbox"/> No. <input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____			

Respectfully submitted,

[Page 1 of 2]

Date May 28, 2004SIGNATURE Cathy A. KodroffREGISTRATION NO. 33,980TYPED or PRINTED NAME Cathy A. Kodroff

(if appropriate)

TELEPHONE 215-540-9200Docket Number: UPN-P3067-P3

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Mail Stop Provisional Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/18 (08-03)

Approved for use through 07/31/2008. OMB 0651-0032
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Docket Number UPN-P3067-P3USA

INVENTOR(S)/APPLICANT(S)		
Given Name (first and middle [if any])	Family or Surname	Residence (City and either State or Foreign Country)
James M.	Wilson	Gladwyne, Pennsylvania

[Page 2 of 2]

Number 1 of 1

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

FEE TRANSMITTAL for FY 2004

Effective 10/01/2003. Patent fees are subject to annual revision.

 Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT (\$ 160.00)

Complete if Known

Application Number	
Filing Date	Herewith
First Named Inventor	Roy et al
Examiner Name	
Art Unit	
Attorney Docket No.	UPN-P3067-P3USA

METHOD OF PAYMENT (check all that apply)

Check Credit card Money Order Other None

 Deposit Account:

Deposit Account Number 08-3040
Deposit Account Name Howson and Howson

The Director is authorized to: (check all that apply)
 Charge fee(s) indicated below Credit any overpayments
 Charge any additional fee(s) or any underpayment of fee(s)
 Charge fee(s) indicated below, except for the filing fee to the above-identified deposit account.

FEE CALCULATION (continued)

3. ADDITIONAL FEES

Large Entity	Small Entity	Fee Description	Fee Paid
1051	130	2051 65 Surcharge - late filing fee or oath	
1052	50	2052 25 Surcharge - late provisional filing fee or cover sheet	
1053	130	1053 130 Non-English specification	
1812	2,520	1812 2,520 For filing a request for ex parte reexamination	
1804	820*	1804 820* Requesting publication of SIR prior to Examiner action	
1805	1,840*	1805 1,840* Requesting publication of SIR after Examiner action	
1251	110	2251 55 Extension for reply within first month	
1252	420	2252 210 Extension for reply within second month	
1253	950	2253 475 Extension for reply within third month	
1254	1,480	2254 740 Extension for reply within fourth month	
1255	2,010	2255 1,005 Extension for reply within fifth month	
1401	330	2401 165 Notice of Appeal	
1402	330	2402 165 Filing a brief in support of an appeal	
1403	290	2403 145 Request for oral hearing	
1451	1,510	1451 1,510 Petition to institute a public use proceeding	
1452	110	2452 55 Petition to revive - unavoidable	
1453	1,330	2453 665 Petition to revive - unintentional	
1501	1,330	2501 665 Utility issue fee (or reissue)	
1502	480	2502 240 Design issue fee	
1503	640	2503 320 Plant issue fee	
1480	130	1460 130 Petitions to the Commissioner	
1807	50	1807 50 Processing fee under 37 CFR 1.17(q)	
1806	180	1806 180 Submission of Information Disclosure Stmt	
8021	40	8021 40 Recording each patent assignment per property (times number of properties)	
1809	770	2809 385 Filing a submission after final rejection (37 CFR 1.129(a))	
1810	770	2810 385 For each additional invention to be examined (37 CFR 1.128(b))	
1801	770	2801 385 Request for Continued Examination (RCE)	
1802	900	1802 900 Request for expedited examination of a design application	

Other fee (specify) _____

*Reduced by Basic Filing Fee Paid

SUBTOTAL (3) (\$)

(\$)

**or number previously paid, if greater; For Reissues, see above

SUBMITTED BY

Name (Print/Type)	Cathy A. Kodroff	Registration No. (Attorney/Agent)	33,980	Telephone	215-540-9200
Signature	<i>Cathy A. Kodroff</i>				
		Date	May 28, 2004		

(Complete if applicable)

WARNING! Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

This collection of information is required by 37 CFR 1.17 and 1.27. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Express Mail No. ER635153476US

CustomerNo. 00270

METHODS OF GENERATING CHIMERIC ADENOVIRUSES AND USES FOR SUCH CHIMERIC ADENOVIRUSES

5 BACKGROUND OF THE INVENTION

The presence of humoral immunity (circulating antibodies) to adenovirus capsid proteins is a barrier to the use of adenovirus vectors for gene therapy. The prototype adenovirus vectors that have been developed for gene therapy are based on subgroup C adenoviruses such as that of serotype 5. The prevalence of neutralizing antibodies against subgroup C adenoviruses is generally high in human populations as a result of frequent exposure to these pathogens. This fact is likely to greatly limit the effectiveness of gene therapy vectors based on serotypes such as Ad5.

Analysis of the nature of the protective antibodies against adenoviruses has indicated that the most important target is the major capsid protein, hexon [Wolfhart 1988] *J. Virol.* 62, 2321; Gall *et al.* (1996) *J. Virol.* 70, 2116]. Several efforts have been made to engineer the hexon so as to evade the anti-hexon antibodies by making chimeric adenoviruses harboring hexons from other serotypes [Roy *et al.* (1998) *J. Virol.* 72, 6875; US Patent No 5922315; Gall *et al.* (1998) *J. Virol.* 72, 10260; Youil *et al.* (2002) *Hum. Gene Ther.* 13, 311; Wu *et al.* (2002) *J. Virol.* 76, 12775].

20 However, this has been largely unsuccessful when exchanges among distant serotypes are attempted.

Alternatively, investigators have proposed using adenovirus vectors that rarely cause human infections or using adenoviruses from non-human sources. However, the lack of a practical manner in which to produce large numbers of such 25 vectors has proved to be a hindrance to developing such vectors.

SUMMARY OF THE INVENTION

The present invention provides a method of modifying adenoviruses having capsids, 30 and particularly, including hexons, from serotypes which are not well adapted for growth in cells useful for adenoviral virion production. The method is useful for

production of scalable amounts of adenoviruses. The modified, or chimeric, adenoviruses are useful for a variety of purposes which are described herein. The invention further provides novel, isolated, adenovirus SA18 nucleic acid and amino acid sequences, vectors containing same, cell lines containing such SA18 sequences and/or vectors, and uses thereof.

Other aspects and advantages of the present invention will be readily apparent from the following detailed description of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 provides the map of the genome of the simian adenovirus generated by shotgun cloning as described in the examples below.

Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus, termed HSS25H5.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides chimeric adenoviruses composed of the left terminal end and right terminal end of an adenovirus which can be cultured in the selected host cell, and the internal regions encoding, at a minimum, the capsid proteins of another adenovirus serotype. This invention is particularly advantageous for generating adenoviruses having serotypes which are difficult to culture in a desired cell type. The invention thus permits generation of chimeric adenoviruses of varying serotypes.

In the embodiments illustrated herein, chimeric adenoviruses have been constructed where most structural proteins, and not merely the hexon or fiber, are derived from an adenovirus of an unrelated serotype, thereby preserving the majority of the protein-protein interactions that are involved in capsid assembly. Most of the early genes such as those encoded by the adenovirus E1 and E4 regions that are responsible for transcription regulation and regulation of the host cell cycle, are retained from a different serotype that is known to result in high titer virus generation in the commonly used cell types, such as HEK 293 which supplies the Ad5 E1 proteins in *trans*.

In another embodiment, the invention provides novel nucleic acid and amino acid sequences from Ad SA18, which was originally isolated from vervet monkey

[ATCC VR-943]. The present invention further provides novel adenovirus vectors and packaging cell lines to produce those vectors for use in the *in vitro* production of recombinant proteins or fragments or other reagents. The invention further provides compositions for use in delivering a heterologous molecule for therapeutic or vaccine purposes. Such therapeutic or vaccine compositions contain the adenoviral vectors carrying an inserted heterologous molecule. In addition, novel sequences of the invention are useful in providing the essential helper functions required for production of recombinant adeno-associated viral (AAV) vectors. Thus, the invention provides helper constructs, methods and cell lines which use these sequences in such production methods.

The term "substantial homology" or "substantial similarity," when referring to a nucleic acid or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 15 99% of the aligned sequences.

The term "substantial homology" or "substantial similarity," when referring to amino acids or fragments thereof, indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid (or its complementary strand), there is amino acid sequence identity in at least about 95 to 20 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, or a fragment thereof which is at least 8 amino acids, or more desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

The term "percent sequence identity" or "identical" in the context of nucleic 25 acid sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome (e.g., about 36 kbp), the full-length of an open reading frame of a gene, protein, subunit, or enzyme [see, e.g., the tables providing the adenoviral coding regions], or a fragment of at least about 500 to 30 5000 nucleotides, is desired. However, identity among smaller fragments, e.g., of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired. Similarly, "percent sequence identity" may be readily determined for amino acid sequences,

over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8 amino acids in length, and may be up to about 700 amino acids. Examples of suitable fragments are described herein.

Identity is readily determined using such algorithms and computer programs as are defined herein at default settings. Preferably, such identity is over the full length of the protein, enzyme, subunit, or over a fragment of at least about 8 amino acids in length. However, identity may be based upon shorter regions, where suited to the use to which the identical gene product is being put.

As described herein, alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs, such as "Clustal W", accessible through Web Servers on the internet. Alternatively, Vector NTI utilities are also used. There are also a number of algorithms known in the art that can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using Fasta, a program in GCG Version 6.1. Fasta provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1, herein incorporated by reference. Similarly programs are available for performing amino acid alignments. Generally, these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program that provides at least the level of identity or alignment as that provided by the referenced algorithms and programs.

As used throughout this specification and the claims, the term "comprise" and its variants including, "comprises", "comprising", among other variants, is inclusive of other components, elements, integers, steps and the like. The term "consists of" or "consisting of" are exclusive of other components, elements, integers, steps and the like.

Except where otherwise specified, the term "vector" includes any genetic element known in the art which will deliver a target molecule to a cell, including, naked DNA, a plasmid, phage, transposon, cosmids, episomes, viruses, etc.

By "minigene" is meant the combination of a selected heterologous gene and the other regulatory elements necessary to drive translation, transcription and/or expression of the gene product in a host cell.

As used herein, the term "transcomplement" refers to when a gene (gene product) of one adenovirus serotype supplies an adenovirus serotype lacking this gene (gene product) from another serotype with the missing function. For example, human adenovirus serotype 5 E1a and E1b functions are known to transcomplement E1-deleted chimpanzee adenovirus Pan 9. Similarly, the inventors have found that human Ad5 E1 transcomplements E1-deleted chimpanzee adenovirus serotypes Pan5, Pan6, Pan7, and simian adenovirus serotypes SV1, SV25 and SV39. Other examples of transcomplementing serotypes include human Ad5 and human Ad2, Ad3, Ad4, Ad5, Ad7, and Ad12.

The term "functionally deleted" or "functional deletion" means that a sufficient amount of the gene region is removed or otherwise damaged, e.g., by mutation or modification, so that the gene region is no longer capable of producing functional products of gene expression. If desired, the entire gene region may be removed. Other suitable sites for gene disruption or deletion are discussed elsewhere in the application.

The term "functional" refers to a product (e.g., a protein or peptide) which performs its native function, although not necessarily at the same level as the native product. The term "functional" may also refer to a gene which encodes and from which a desired product can be expressed.

I. Chimeric Adenoviral Vectors

The compositions of this invention include chimeric adenoviral vectors that deliver a heterologous molecule to cells. For delivery of such a heterologous molecule, the vector can be a plasmid or, preferably, a chimeric adenovirus. The chimeric adenoviruses of the invention include adenovirus DNA from at least two source serotypes, a "donating serotype" and a "parental adenovirus" as described in more detail herein, and a minigene.

Because the adenoviral genome contains open reading frames on both strands, in many instances reference is made herein to 5' and 3' ends of the various regions to avoid confusion between specific open reading frames and gene regions. Thus, when

reference is made herein to the "left" and "right" end of the adenoviral genome, this reference is to the ends of the approximately 36 kb adenoviral genome when depicted in schematic form as is conventional in the art [see, e.g., Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)]. Thus, as used
5 herein, the "left terminal end" of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme left end of the schematic. Typically, the left end refers to be portion of the genome beginning at map unit 0 and extending to the right to include at least the 5' inverted terminal repeats (ITRs), and excludes the internal regions of the
10 genome encoding the structural genes. As used herein, the "right terminal end" of the adenoviral genome refers to portion of the adenoviral genome which, when the genome is depicted schematically in linear form, is located at the extreme right end of the schematic. Typically, the right end of the adenoviral genome refers to be portion of the genome ending at map unit 36 and extending to the left to include at least the 3'
15 ITRs, and excludes the internal regions of the genome encoding the structural genes.

A. Adenovirus Regulatory Sequences

1. Serotype

The selection of the adenovirus serotype donating its left terminal end and right terminal end can be readily made by one of skill in the art from
20 among serotypes which can readily be cultured in the desired cell line. Among other factors which may be considered in selecting the serotype of the donating serotype is compatibility with the adenovirus serotype which will be supplying the internal regions at the location at which their sequences are hybridized.

Suitable adenoviruses for donating their left and right termini
25 are available from the American Type Culture Collection, Manassas, Virginia, US (ATCC), a variety of academic and commercial sources, or the desired regions of the donating adenoviruses may be synthesized using known techniques with reference to sequences published in the literature or available from databases (e.g., GenBank, etc.). Examples of suitable donating adenoviruses include, without limitation, human
30 adenovirus serotypes 2, 3, 4, 5, 7, and 12, and further including any of the presently identified human types [see, e.g., Horwitz, "Adenoviridae and Their Replication", in *VIROLOGY*, 2d ed., pp. 1679-1721 (1990)] which can be cultured in the desired cell. Similarly adenoviruses known to infect non-human primates (e.g., chimpanzees,

rhesus, macaque, and other simian species) or other non-human mammals and which grow in the desired cell can be employed in the vector constructs of this invention.

Such serotypes include, without limitation, chimpanzee adenoviruses Pan 5 [VR-591], Pan6 [VR-592], Pan7 [VR-593], and C68 (Pan9), described in US Patent No.

5 6,083,716; and simian adenoviruses including, without limitation SV1 [VR-195];
SV25 [SV-201]; SV35; SV15; SV-34; SV-36; SV-37, and baboon adenovirus [VR-
275], among others. The sequences of Pan 5 (also termed C5), Pan 6 (also termed
C6), Pan 7 (also termed C7), SV1, SV25, and SV39 have been described [WO
03/046124, published 5 June 2003; and in US Patent Application No. 10/739,096,
10 filed December 19, 2003)], which are incorporated by reference. In the following
examples, the human 293 cells and adenovirus type 5 (Ad5), Pan9, and Ad40 are used
for convenience. However, one of skill in the art will understand that other cell lines
and/or comparable regions derived from other adenoviral strains may be readily
selected and used in the present invention in the place of (or in combination with)
15 these serotypes.

2. Sequences

The minimum sequences which must be supplied by the adenovirus donating its left terminal end and its right terminal end include the 5' cis-elements and the 3' cis-elements necessary for replication and packaging. Typically, 20 the 5' cis-elements necessary for packaging and replication include the 5' inverted terminal repeat (ITR) sequences (which functions as origins of replication) and the native 5' packaging enhancer domains (that contain sequences necessary for packaging linear Ad genomes and enhancer elements for the E1 promoter). The right end of the adenoviral genome includes the 3' cis-elements (including the ITRs) 25 necessary for packaging and encapsidation. Desirably, the adenovirus serotype donating its left and right termini and/or an adenovirus serotype which transcomplements the serotype of the donating adenovirus, further provides the functions of the necessary adenovirus early genes, including E1 (E1a and E1b), E2 (E2a and E2b), and E4 (including at least the ORF6 region). E3 is not essential and 30 may be deleted as desired, e.g., for insertion of a transgene in this region or to provide space for a transgene inserted in another region (typically for packaging it is desirable for the total adenoviral genome to be under 36 kb).

In certain embodiments, the necessary adenovirus early genes are contained in the chimeric construct of the invention. In other embodiment, one or more of the necessary adenovirus early genes can be provided by the packaging host cell or in *trans*.

5 In general, the chimeric adenovirus of the invention contains regulatory sequences from the donating adenovirus serotype, or a transcomplementing serotype, to provide the chimeric adenovirus with compatible regulatory proteins. Optionally, one or more of the necessary adenoviral structural genes is provided by the adenovirus donating its left terminal and its right terminal end.

10 In certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the adenovirus serotype donating its left and right termini. Where all of these genes are

15 derived from the adenovirus serotype donating the 5' and 3' ITRs, a "pseudotyped" chimeric is formed. In one embodiment, the chimeric adenovirus contains the left end of the adenovirus genome from the donating serotype, from the 5' ITR through the end of the pol gene (or the pTP). In another embodiment, the chimeric adenovirus contains the left end of the donating adenovirus serotype, from the 5' ITR through the penton. In yet another embodiment, the chimeric adenovirus contains the left end of the donating adenovirus serotype, e.g., through the end of pTP, but contains an ITR from an adenovirus serotype heterologous to the donating adenovirus serotype. Still other embodiments will be readily apparent from the present disclosure.

20 Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins (e.g., without limitation, polymerase, terminal protein, IIIa protein). Suitably, these genes express functional proteins which permit packaging of the adenovirus genes into the capsid. Alternatively, one or more of these proteins (whether hybrid or non-hybrid) can be functionally deleted in the chimeric adenovirus. Where desired, any necessary proteins functionally deleted in the chimeric adenovirus can be expressed in *trans* in the packaging cell.

B. Parental Adenovirus Structural Proteins

1. Serotypes

This invention is particularly well adapted for use in generating chimeric adenoviruses in which the capsid proteins are from a parental adenovirus which does not efficiently grow in a desirable host cell. The selection of the parental adenovirus serotype providing the internal regions can be readily made by one of skill in the art based on the information provided herein.

A variety of suitable adenoviruses can serve as a parental adenovirus supplying the regions encoding the structural (*i.e.*, capsid proteins). Many such adenoviruses can be obtained from the same sources as described above for the donating adenovirus serotypes. Examples of suitable parental adenovirus serotypes includes, without limitation, human adenovirus serotype 40, among others [see, e.g., Horwitz, "Adenoviridae and Their Replication", in VIROLOGY, 2d ed., pp. 1679-1721 (1990)], and adenoviruses known to infect non-human primates (*e.g.*, chimpanzees, rhesus, macaque, and other simian species) or other non-human mammals, including, without limitation, chimpanzee adenovirus C1, described in US Patent No. 6,083,716; simian adenoviruses, and baboon adenoviruses, among others. In addition, the parental adenovirus supplying the internal regions may be from a non-naturally occurring adenovirus serotype, such as may be generated using a variety of techniques known to those of skill in the art.

In one embodiment illustrated herein, a chimeric virus that was constructed was that between the chimpanzee adenoviruses Pan-5 and C1 exhibited a higher titer in human 293 cells than the wild-type parental virus. However, the invention is not limited to the use of these chimpanzee adenoviruses, or to the combination of simian-simian, human-human, or simian-human chimeric adenoviruses. For example, it may be desirable to utilize bovine or canine adenoviruses, or other non-human mammalian adenoviruses which do not naturally infect and/or replicate in human cells.

In the following examples, the human adenovirus type 40 (Ad40) and the chimpanzee adenovirus C1, simian Pan 5 and Ad40, and Pan 5 and simian adenovirus SA18, are used. However, one of skill in the art will understand that other adenoviral serotypes may be readily selected and used in the present invention in the place of (or in combination with) these serotypes.

2. Sequences

The parental adenovirus provides to the chimeric construct of the invention its internal regions which includes structural proteins necessary for generating a capsid having the desired characteristics of the parental adenovirus.

5 These desired characteristics include, but are not limited to, the ability to infect target cells and delivery a heterologous transgene, the ability to elude neutralizing antibodies directed to another adenovirus serotype (i.e., avoiding clearance due to cross-reactivity), and/or the ability to infect cells in the absence of an immune response to the chimeric adenovirus. The advantages of such characteristics may be
10 most readily apparent in a regimen which involves repeat delivery of adenoviral vectors. The left and right termini of the parent adenovirus, including at least the 5' ITRs, the E1 region, the E4 region and the 3' ITRs are non-functional and, preferably, completely absent. Optionally, all adenovirus regulatory proteins from this parental adenovirus are non-functional and only the structural proteins (or selected structural
15 proteins) are retained.

At a minimum, the parental adenovirus provides the adenoviral late region encoding the hexon protein. Suitably, the parental adenovirus further provides the late regions encoding the penton and the fiber. In certain embodiments, all of the functional adenoviral late regions, including L1 (encoding
20 52/55 Da, IIIa proteins), L2 (encoding penton, VII, V, Mu proteins), L3 (encoding VI, hexon, Endoprotease), L4 (encoding 100 kD, 33 kD, VIII proteins) and L5 (encoding fiber protein) are supplied by the parental adenovirus. Optionally, one or more of these late gene functions, with the exception of those encoding the hexon, penton and fiber proteins, can be functionally deleted. Any necessary structural proteins may be
25 supplied in *trans*.

Thus, in certain embodiments, the chimeric adenovirus further contains one or more functional adenovirus genes, including, the Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and/or protein VI from the
30 parental adenovirus donating its internal regions. Optionally, one or more of the genes can be hybrids formed from the fusion of the donating adenovirus serotype and the parental adenovirus serotype providing the capsid proteins, as described above.

C. The "Minigene"

Typically, an adenoviral vector of the invention is designed to contain a minigene which may be inserted into the site of a partially deleted, fully deleted (absent), or disrupted adenoviral gene. For example, the minigene may be located in 5 the site of such a functional E1 deletion or functional E3 deletion, or another suitable site.

The methods employed for the selection of the transgene, the cloning and construction of the "minigene" and its insertion into the viral vector are within the skill in the art given the teachings provided herein.

10 1. The transgene

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a polypeptide, protein, or other product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation, 15 and/or expression in a host cell.

The composition of the transgene sequence will depend upon the use to which the adenoviral vector will be put. For example, the adenoviral vector may be used as a helper virus in production of recombinant adeno-associated viruses or in production of recombinant adenoviruses deleted of essential adenoviral gene 20 functions which are supplied by the adenoviral vector, or for a variety of production uses. Alternatively, the adenoviral vector may be used for diagnostic purposes.

One type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -lactamase, β -galactosidase 25 (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane 30 bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc. These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other

spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for beta-galactosidase activity. Where the transgene is GFP or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins,

10 peptides, RNA, enzymes, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, si RNAs, small hairpin RNAs, trans-splicing RNAs, catalytic RNAs, and antisense RNAs. One example of a useful RNA sequence is a sequence which extinguishes expression of a targeted nucleic acid sequence in the treated animal.

15 The transgene may be used for treatment, e.g., of genetic deficiencies, as a cancer therapeutic or vaccine, for induction of an immune response, and/or for prophylactic vaccine purposes. As used herein, induction of an immune response refers to the ability of a molecule (e.g., a gene product) to induce a T cell and/or a humoral immune response to the molecule. The invention further includes

20 using multiple transgenes, e.g., to correct or ameliorate a condition caused by a multi-subunit protein. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, e.g., for an immunoglobulin, the platelet-derived growth factor, or a dystrophin protein. In order

25 for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when

30 the size of the DNA encoding each of the subunits is small, e.g., the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, e.g., M.L. Donnelly, *et al*, *J. Gen.*

Virol., 78(Pt 1):13-21 (Jan 1997); Furler, S., et al, *Gene Ther.*, 8(11):864-873 (June 2001); Klump H., et al., *Gene Ther.*, 8(10):811-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. However, the selected transgene may encode any biologically active product or other product, e.g., a product desirable for study.

5 Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

2. Vector and Transgene Regulatory Elements

10 In addition to the major elements identified above for the minigene, the adenoviral vector also includes conventional control elements which are operably linked to the transgene in a manner that permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked"

15 sequences include both expression control sequences that are contiguous with the gene of interest and expression control sequences that act in trans or at a distance to control the gene of interest.

20 Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (i.e., Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters which are native, constitutive, 25 inducible and/or tissue-specific, are known in the art and may be utilized.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, e.g., Boshart et al, *Cell*, 41:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β-actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1α promoter [Invitrogen].

30 Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as

temperature, or the presence of a specific physiological state, e.g., acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems
5 have been described and can be readily selected by one of skill in the art. For example, inducible promoters include the zinc-inducible sheep metallothioneine (MT) promoter and the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter. Other inducible systems include the T7 polymerase promoter system [WO 98/10088]; the ecdysone insect promoter [No *et al*, *Proc. Natl. Acad. Sci. USA*, 93:3346-3351 (1996)], the tetracycline-repressible system [Gossen *et al*,
10 *Proc. Natl. Acad. Sci. USA*, 89:5547-5551 (1992)], the tetracycline-inducible system [Gossen *et al*, *Science*, 268:1766-1769 (1995), see also Harvey *et al*, *Curr. Opin. Chem. Biol.*, 2:512-518 (1998)]. Other systems include the FK506 dimer, VP16 or p65 using castradiol, diphenol murslerone, the RU486-inducible system [Wang *et al*,
15 *Nat. Biotech.*, 15:239-243 (1997) and Wang *et al*, *Gene Ther.*, 4:432-441 (1997)] and the rapamycin-inducible system [Magari *et al*, *J. Clin. Invest.*, 100:2865-2872 (1997)]. The effectiveness of some inducible promoters increases over time. In such cases one can enhance the effectiveness of such systems by inserting multiple repressors in tandem, e.g., TetR linked to a TetR by an IRES. Alternatively, one can
20 wait at least 3 days before screening for the desired function. One can enhance expression of desired proteins by known means to enhance the effectiveness of this system. For example, using the Woodchuck Hepatitis Virus Posttranscriptional Regulatory Element (WPRE).

In another embodiment, the native promoter for the transgene
25 will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.
30

Another embodiment of the transgene includes a transgene operably linked to a tissue-specific promoter. For instance, if expression in skeletal

muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β-actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally occurring promoters (see Li *et al.*, *Nat. Biotech.*, 17:241-245 (1999)).

5 Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake *et al.*, *J. Virol.*, 71:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al.*, *Gene Ther.*, 3:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot *et al.*, *Hum. Gene Ther.*, 7:1503-14 (1996)), bone osteocalcin (Stein *et al.*, *Mol. Biol. Rep.*, 24:185-96 (1997)); bone sialoprotein (Chen *et al.*, *J. Bone Miner. Res.*, 11:654-64

10 10 (1996)), lymphocytes (CD2, Hansal *et al.*, *J. Immunol.*, 161:1063-8 (1998); immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen *et al.*, *Cell. Mol. Neurobiol.*, 13:503-15 (1993)), neurofilament light-chain gene (Piccioli *et al.*, *Proc. Natl. Acad. Sci. USA*, 88:5611-5 (1991)), and the neuron-specific vgf gene (Piccioli *et al.*, *Neuron*, 15:373-84 (1995)),

15 among others.

Optionally, vectors carrying transgenes encoding therapeutically useful or immunogenic products may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be packaged into a viral particle) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the vector may include an origin of replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, e.g., Sambrook *et al.*, and references cited therein].

These vectors are generated using the techniques and sequences provided herein, in conjunction with techniques known to those of skill in the art. Such techniques include conventional cloning techniques of cDNA such as those described in texts [Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 30 Cold Spring Harbor Press, Cold Spring Harbor, NY], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence.

II. Production of the Recombinant Viral Particle

In one embodiment, the invention provides a method of generating recombinant chimeric adenoviral particles in which the capsid of the chimeric adenovirus is of a serotype incapable of efficient growth in the selected host cell. A vector suitable for production of recombinant chimeric adenoviral particles can be generated by direct cloning. Alternatively, such particles can be generated by homologous recombination between a first vector containing the left end of the chimeric adenoviral genome and a second vector containing the right end of the chimeric adenoviral genome. However, any suitable methodology known to those of skill in the art can be readily utilized to generate a vector suitable to generate a production vector, preferably which contains the entire chimeric adenoviral genome, including a minigene. This production vector is then introduced into a host cell in which the adenoviral capsid protein is assembled and the chimeric adenoviral particle assembled as described.

The chimeric adenoviruses of the invention include those in which one or more adenoviral genes are absent, or otherwise rendered non-functional. If any of the missing gene functions are essential to the replication and infectivity of the adenoviral particle, these functions are supplied by a complementation (or transcomplementing) cell line or a helper vector expressing these functions during production of the chimeric adenoviral particle.

Examples of chimeric adenoviruses containing such missing adenoviral gene functions include those which are partially or completely deleted in the E1a and/or E1b gene. In such a case, the E1 gene functions can be supplied by the packaging host cell, permitting the chimeric construct to be deleted of E1 gene functions and, if desired, for a transgene to be inserted in this region. Optionally, the E1 gene can be of a serotype which transcomplements the serotype providing the other adenovirus sequences in order to further reduce the possibility of recombination and improve safety. In other embodiments, it is desirable to retain an intact E1a and/or E1b region in the recombinant adenoviruses. Such an intact E1 region may be located in its native location in the adenoviral genome or placed in the site of a deletion in the native adenoviral genome (e.g., in the E3 region).

In another example, all or a portion of the adenovirus delayed early gene E3 may be eliminated from the chimeric adenovirus. The function of adenovirus E3 is

believed to be irrelevant to the function and production of the recombinant virus particle. Chimeric adenovirus vectors may also be constructed having a deletion of at least the ORF6 region of the E4 gene, and more desirably because of the redundancy in the function of this region, the entire E4 region. Still another vector of this invention contains a deletion in the delayed early gene E2a. Similarly, deletions in the intermediate genes IX and IVa₂ may be useful for some purposes. Optionally, deletions may also be made in selected portions of the late genes L1 through L5, as described above.

Other deletions may be made in the other structural or non-structural 5 adenovirus genes. The above-discussed deletions may be used individually, i.e., an adenovirus sequence for use in the present invention may contain deletions in only a single region. Alternatively, deletions of entire genes or portions thereof effective to destroy their biological activity may be used in any combination. For example, in one exemplary vector, the adenovirus sequence may have deletions of the E1 genes and 10 the E4 gene, or of the E1, E2a and E3 genes, or of the E1 and E3 genes, or of E1, E2a and E4 genes, with or without deletion of E3, and so on. As discussed above, such 15 deletions may be used in combination with other mutations, such as temperature-sensitive mutations, to achieve a desired result.

Examples of suitable transcomplementing serotypes are provided above. The 20 use of transcomplementing serotypes can be particularly advantageous where there is diversity between the Ad sequences in the vector of the invention and the human AdE1 sequences found in currently available packaging cells. In such cases, the use of the current human E1-containing cells prevents the generation of replication-competent adenoviruses during the replication and production process. However, in 25 certain circumstances, it will be desirable to utilize a cell line which expresses the E1 gene products can be utilized for production of an E1-deleted simian adenovirus. Such cell lines have been described. See, e.g., US Patent 6,083,716.

A. Packaging Host Cells

Suitably, the packaging host cell is selected from among cells in 30 which the adenovirus serotype donating the left and right terminal ends of the chimeric genome are capable of efficient growth. The host cells are preferably of mammalian origin, and most preferably are of non-human primate or human origin.

Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549 [ATCC Accession No. CCL 185], 911 cells, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, FJ *et al*, (1998), *Hum Gene Ther*, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2, HeLa [ATCC Accession No. CCL 2], KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. These cell lines are all available from the American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209. Other suitable cell lines may be obtained from other sources. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc.

As described above, a chimeric adenovirus of the invention can lack one or more functional adenoviral regulatory and/or structural genes which are supplied either by the host cell or in *trans* to effect packaging of the chimeric adenovirus into the viral capsid to generate the viral particle. Thus, the ability of a selected host cell to supply transcomplementing adenoviral sequences may be taken into consideration in selecting a desired host cell.

In one example, the cells are from a stable cell line which expresses adenovirus E1a and E1b functions from a cell line which transcomplements the adenovirus serotype which donates the left and right termini to the chimera of the invention, permitting the chimera to be E1-deleted. Alternatively, where the cell line does not transcomplement the adenovirus donating the termini, E1 functions may be provided by the chimera, or in *trans*.

If desired, one may utilize the sequences provided herein to generate a packaging cell or cell line that expresses, at a minimum, the adenovirus E1 gene from the adenovirus serotype donating the 5' ITR under the transcriptional control of a promoter for expression, or a transcomplementing serotype, in a selected parent cell line. Inducible or constitutive promoters may be employed for this purpose. Examples of such promoters are described in detail elsewhere in this specification. A parent cell is selected for the generation of a novel cell line expressing any desired adenovirus or adenovirus gene, including, e.g., a human Ad5, AdPan5, Pan6, Pan7,

SV1, SV25 or SV39 gene. Without limitation, such a parent cell line may be HeLa [ATCC Accession No. CCL 2], A549 [ATCC Accession No. CCL 185], HEK 293, KB [CCL 17], Detroit [e.g., Detroit 510, CCL 72] and WI-38 [CCL 75] cells, among others. Many of these cell lines are all available from the ATCC. Other suitable

5 parent cell lines may be obtained from other sources.

Such E1-expressing cell lines are useful in the generation of chimeric adenovirus E1 deleted vectors. Additionally, or alternatively, the invention provides cell lines that express one or more simian adenoviral gene products, e.g., E1a, E1b, E2a, and/or E4 ORF6, can be constructed using essentially the same procedures for
10 use in the generation of chimeric viral vectors. Such cell lines can be utilized to transcomplement adenovirus vectors deleted in the essential genes that encode those products, or to provide helper functions necessary for packaging of a helper-dependent virus (e.g., adeno-associated virus). The preparation of a host cell according to this invention involves techniques such as assembly of selected DNA
15 sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook et al., cited above, use of overlapping oligonucleotide sequences of the adenovirus genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired
20 nucleotide sequence.

In still another alternative, the essential adenoviral gene products are provided in *trans* by the adenoviral vector and/or helper virus. In such an instance, a suitable host cell can be selected from any biological organism, including prokaryotic (e.g., bacterial) cells, and eukaryotic cells, including, insect cells, yeast cells and
25 mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, HEK 293 cells or PERC6 (both of which express functional adenoviral E1) [Fallaux, FJ et al. (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from
30 mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, i.e., fibroblast, hepatocyte, tumor cell, etc.

B. Helper Vectors

Thus, depending upon the adenovirus gene content of the adenoviral vectors and any adenoviral gene functions expressed from the host cell, a helper vector may be necessary to provide sufficient adenovirus gene sequences necessary to produce an infective recombinant viral particle containing the minigene. See, for example, the techniques described for preparation of a "minimal" human Ad vector in International Patent Application WO96/13597, published May 9, 1996, and incorporated herein by reference. Suitably, these helper vectors may be non-replicating genetic elements, a plasmid, or a virus.

Useful helper vectors contain selected adenovirus gene sequences not present in the adenovirus vector construct and/or not expressed by the packaging cell line in which the vector is transfected. In one embodiment, the helper virus is replication-defective and contains a variety of adenovirus genes in addition to the sequences described above. Such a helper vector is desirably used in combination with an E1-expressing cell line.

Helper vectors may be formed into poly-cation conjugates as described in Wu *et al*, *J. Biol. Chem.*, 264:16985-16987 (1989); K. J. Fisher and J. M. Wilson, *Biochem. J.*, 299:49 (April 1, 1994). A helper vector may optionally contain a second reporter minigene. A number of such reporter genes are known to the art. The presence of a reporter gene on the helper virus which is different from the transgene on the adenovirus vector allows both the Ad vector and the helper vector to be independently monitored. This second reporter is used to enable separation between the resulting recombinant virus and the helper virus upon purification.

C. Assembly of Viral Particle and Transfection of a Cell Line

Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 μ g to about 100 μ g DNA, and preferably about 10 to about 50 μ g DNA to about 1×10^4 cells to about 1×10^{13} cells, and preferably about 10^5 cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed above, including transfection, and infection. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed transiently. The gene products

may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently.

Furthermore, the promoters for each of the adenoviral genes may be selected independently from a constitutive promoter, an inducible promoter or a native adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (i.e., by the differentiation state or in replicating or quiescent cells) or by exogenously added factors, for example.

Introduction of the molecules (as plasmids or viruses) into the host cell may also be accomplished using techniques known to the skilled artisan and as discussed throughout the specification. In preferred embodiment, standard transfection techniques are used, e.g., CaPO₄ transfection or electroporation.

Assembly of the selected DNA sequences of the adenovirus (as well as the transgene and other vector elements) into various intermediate plasmids, and the use of the plasmids and vectors to produce a recombinant viral particle are all achieved using conventional techniques. Such techniques include direct cloning as described [G. Gao *et al*, *Gene Ther.* 2003 Oct; 10(22):1926-1930; US Patent Publication No. 2003-0092161-A, published May 15, 2003; International Patent Application No. PCT/US03/12405]. Other cloning techniques of cDNA such as those described in texts [Sambrook *et al*, cited above], use of overlapping oligonucleotide sequences of the adenovirus genomes, polymerase chain reaction, and any suitable method which provides the desired nucleotide sequence can be utilized. Standard transfection and co-transfection techniques are employed, e.g., CaPO₄ precipitation techniques. Other conventional methods employed include homologous recombination of the viral genomes, plaquing of viruses in agar overlay, methods of measuring signal generation, and the like.

For example, following the construction and assembly of the desired minigene-containing viral vector, the vector is transfected in vitro in the presence of an optional helper vector into the packaging cell line. The functions expressed from the plasmid, packaging cell line and helper virus, if any, permits the adenovirus-transgene sequences in the vector to be replicated and packaged into virion capsids, resulting in the chimeric viral particles. The current method for producing such virus particles is transfection-based. However, the invention is not limited to such

methods. The resulting chimeric adenoviruses are useful in transferring a selected transgene to a selected cell.

III. Use of the Chimeric Adenovirus Vectors

5 The chimeric adenovirus vectors of the invention are useful for gene transfer to a human or veterinary subject (including, non-human primates, non-simian primates, and other mammals) *in vitro*, *ex vivo*, and *in vivo*.

10 The recombinant adenovirus vectors described herein can be used as expression vectors for the production of the products encoded by the heterologous genes *in vitro*. For example, the recombinant adenoviruses containing a gene inserted into the location of an E1 deletion may be transfected into an E1-expressing cell line as described above. Alternatively, replication-competent adenoviruses may be used in another selected cell line. The transfected cells are then cultured in the conventional manner, allowing the recombinant adenovirus to express the gene
15 product from the promoter. The gene product may then be recovered from the culture medium by known conventional methods of protein isolation and recovery from culture.

20 A chimeric adenoviral vector of the invention provides an efficient gene transfer vehicle that can deliver a selected transgene to a selected host cell *in vivo* or *ex vivo* even where the organism has neutralizing antibodies to one or more AAV serotypes. In one embodiment, the rAd and the cells are mixed *ex vivo*; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient. These compositions are particularly well suited to gene delivery for therapeutic purposes and for immunization, including inducing protective immunity.
25

30 More commonly, the chimeric adenoviral vectors of the invention will be utilized for delivery of therapeutic or immunogenic molecules, as described below. It will be readily understood for both applications that the recombinant adenoviral vectors of the invention are particularly well suited for use in regimens involving repeat delivery of recombinant adenoviral vectors. Such regimens typically involve delivery of a series of viral vectors in which the viral capsids are alternated. The viral capsids may be changed for each subsequent administration, or after a pre-selected number of administrations of a particular serotype capsid (e.g., one, two, three, four

or more). Thus, a regimen may involve delivery of a rAd with a first capsid, delivery with a rAd with a second capsid, and delivery with a third capsid. A variety of other regimens which use the Ad capsids of the invention alone, in combination with one another, or in combination with other Ad serotypes will be apparent to those of skill in the art. Optionally, such a regimen may involve administration of rAd with capsids of non-human primate adenoviruses, human adenoviruses, or artificial (e.g., chimeric) serotypes such as are described herein. Each phase of the regimen may involve administration of a series of injections (or other delivery routes) with a single Ad serotype capsid followed by a series with another Ad serotype capsid. Alternatively, the recombinant Ad vectors of the invention may be utilized in regimens involving other non-adenoviral-mediated delivery systems, including other viral systems, non-viral delivery systems, protein, peptides, and other biologically active molecules.

The following sections will focus on exemplary molecules which may be delivered via the adenoviral vectors of the invention.

A. Ad-Mediated Delivery of Therapeutic Molecules

In one embodiment, the Ad vectors described herein are administered to humans according to published methods for gene therapy. A viral vector of the invention bearing the selected transgene may be administered to a patient, preferably suspended in a biologically compatible solution or pharmaceutically acceptable delivery vehicle. A suitable vehicle includes sterile saline. Other aqueous and non-aqueous isotonic sterile injection solutions and aqueous and non-aqueous sterile suspensions known to be pharmaceutically acceptable carriers and well known to those of skill in the art may be employed for this purpose.

The adenoviral vectors are administered in sufficient amounts to transduce the target cells and to provide sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to the retina and other intraocular delivery methods, direct delivery to the liver, inhalation, intranasal, intravenous, intramuscular, intratracheal, subcutaneous, intradermal, rectal, oral and other parenteral routes of administration. Routes of administration may be combined, if

desired, or adjusted depending upon the transgene or the condition. The route of administration primarily will depend on the nature of the condition being treated.

Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective adult human or veterinary dosage of the viral vector is generally in the range of from about 100 μ L to about 100 mL of a carrier containing concentrations of from about 1×10^6 to about 1×10^{15} particles, about 1×10^{11} to 1×10^{13} particles, or about 1×10^9 to 1×10^{12} particles. Dosages will range depending upon the size of the animal and the route of administration. For example, a suitable human or veterinary dosage (for about an 80 kg animal) for intramuscular injection is in the range of about 1×10^9 to about 5×10^{12} particles per mL, for a single site. Optionally, multiple sites of administration may be delivered. In another example, a suitable human or veterinary dosage may be in the range of about 1×10^{11} to about 1×10^{15} particles for an oral formulation. One of skill in the art may adjust these doses, depending the route of administration, and the therapeutic or vaccinal application for which the recombinant vector is employed. The levels of expression of the transgene, or for an immunogen, the level of circulating antibody, can be monitored to determine the frequency of dosage administration. Yet other methods for determining the timing of frequency of administration will be readily apparent to one of skill in the art.

An optional method step involves the co-administration to the patient, either concurrently with, or before or after administration of the viral vector, of a suitable amount of a short acting immune modulator. The selected immune modulator is defined herein as an agent capable of inhibiting the formation of neutralizing antibodies directed against the recombinant vector of this invention or capable of inhibiting cytolytic T lymphocyte (CTL) elimination of the vector. The immune modulator may interfere with the interactions between the T helper subsets (T_{H1} or T_{H2}) and B cells to inhibit neutralizing antibody formation. Alternatively, the immune modulator may inhibit the interaction between T_{H1} cells and CTLs to reduce the occurrence of CTL elimination of the vector. A variety of useful immune modulators and dosages for use of same are disclosed, for example, in Yang *et al.*, *J. Virol.*, 70(9) (Sept 1996); International Patent Application No. WO96/12406, published May 2, 1996; and International Patent Application No.PCT/US96/03035,

all incorporated herein by reference. Typically, such immune modulators would be selected when the transgene is a therapeutic which requires repeat delivery.

1. Therapeutic Transgenes

Useful therapeutic products encoded by the transgene include

- 5 hormones and growth and differentiation factors including, without limitation, insulin, glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF),
- 10 erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), transforming growth factor α (TGF α), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-II), any one of the transforming growth factor superfamily, including TGF, activins, inhibins, or any of
- 15 the bone morphogenic proteins (BMP) BMPs 1-15, any one of the heregluin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of
- 20 semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, e.g., IL-2, IL-4, IL-12 and IL-18), monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors and, interferons, and, stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without limitation, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins, humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include

complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CF2 and CD59.

Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and

5 immune system proteins. The invention encompasses receptors for cholesterol regulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, proteins useful in the regulation of lipids, including, e.g., apolipoprotein (apo) A and its isoforms (e.g., ApoAI), apoE and its isoforms including E2, E3 and E4), SRB1, 10 ABC1, and the scavenger receptor. The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as jun, fos, max, mad, serum response factor (SRF), AP-1, AP2, myb, MyoD and 15 myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding protein, STAT, GATA-box binding proteins, e.g., GATA-3, and the forkhead family of winged helix proteins.

20 Other useful gene products include, carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, fumarylacetate hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, porphobilinogen deaminase, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-CoA dehydrogenase,

25 propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylate, hepatic phosphorylase, phosphorylase kinase, glycine decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin cDNA sequence. Other useful gene products include those useful for treatment of

30 hemophilia A (e.g., Factor VIII and its variants, including the light chain and heavy chain of the heterodimer, optionally operably linked by a junction), and the B-domain deleted Factor VIII, see US 6,200,560 and 6,221,349], and useful for treatment of hemophilia B (e.g, Factor IX).

Still other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful in certain immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

Reduction and/or modulation of expression of a gene are particularly desirable for treatment of hyperproliferative conditions characterized by

5 hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGRF. In addition to oncogene products as

10 target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found

15 at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

20

Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that

25 are associated with autoimmunity including cell receptors and cells which produce self-directed antibodies. T-cell mediated autoimmune diseases include rheumatoid arthritis (RA), multiple sclerosis (MS), Sjögren's syndrome, sarcoidosis, insulin dependent diabetes mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis,

30 vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

The chimeric adenoviral vectors of the invention are particularly well suited for therapeutic regimens in which multiple adenoviral-mediated deliveries of transgenes is desired, e.g., in regimens involving redelivery of the same transgene or in combination regimens involving delivery of other transgenes. Such regimens may involve administration of a chimeric adenoviral vector, followed by re-administration with a vector from the same serotype adenovirus. Particularly desirable regimens involve administration of a chimeric adenoviral vector of the invention, in which the serotype of the viral vector delivered in the first administration differs from the serotype of the viral vector utilized in one or more of the subsequent administrations.

For example, a therapeutic regimen involves administration of a chimeric vector and repeat administration with one or more adenoviral vectors of the same or different serotypes. In another example, a therapeutic regimen involves administration of an adenoviral vector followed by repeat administration with a chimeric vector of the invention which differs from the serotype of the first delivered adenoviral vector, and optionally further administration with another vector which is the same or, preferably, differs from the serotype of the vector in the prior administration steps. These regimens are not limited to delivery of adenoviral vectors constructed using the chimeric serotypes of the invention. Rather, these regimens can readily utilize chimeric or non-chimeric vectors of other adenoviral serotypes, which may be of artificial, human or non-human primate, or other mammalian sources, in combination with one or more of the chimeric vectors of the invention. Examples of such serotypes are discussed elsewhere in this document. Further, these therapeutic regimens may involve either simultaneous or sequential delivery of chimeric adenoviral vectors of the invention in combination with non-adenoviral vectors, non-viral vectors, and/or a variety of other therapeutically useful compounds or molecules. The present invention is not limited to these therapeutic regimens, a variety of which will be readily apparent to one of skill in the art.

B. Ad-Mediated Delivery of Immunogenic Transgenes

The adenoviruses of the invention may also be employed as immunogenic compositions. As used herein, an immunogenic composition is a composition to which a humoral (e.g., antibody) or cellular (e.g., a cytotoxic T cell) response is mounted to a transgene product delivered by the immunogenic composition following delivery to a mammal, and preferably a primate. The present

invention provides an Ad that can contain in any of its adenovirus sequence deletions a gene encoding a desired immunogen. Chimeric adenoviruses based on simian or other non-human mammalian primate serotypes are likely to be better suited for use as a live recombinant virus vaccine in different animal species compared to an adenovirus of human origin, but is not limited to such a use. The recombinant adenoviruses can be used as prophylactic or therapeutic vaccines against any pathogen for which the antigen(s) crucial for induction of an immune response and able to limit the spread of the pathogen has been identified and for which the cDNA is available.

Such vaccinal (or other immunogenic) compositions are formulated in a suitable delivery vehicle, as described above. Generally, doses for the immunogenic compositions are in the range defined above for therapeutic compositions. The levels of immunity of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of antibody titers in the serum, optional booster immunizations may be desired.

Optionally, a vaccinal composition of the invention may be formulated to contain other components, including, e.g. adjuvants, stabilizers, pH adjusters, preservatives and the like. Such components are well known to those of skill in the vaccine art. Examples of suitable adjuvants include, without limitation, liposomes, alum, monophosphoryl lipid A, and any biologically active factor, such as cytokine, an interleukin, a chemokine, a ligands, and optimally combinations thereof. Certain of these biologically active factors can be expressed in vivo, e.g., via a plasmid or viral vector. For example, such an adjuvant can be administered with a priming DNA vaccine encoding an antigen to enhance the antigen-specific immune response compared with the immune response generated upon priming with a DNA vaccine encoding the antigen only.

The adenoviruses are administered in "an immunogenic amount", that is, an amount of adenovirus that is effective in a route of administration to transfect the desired cells and provide sufficient levels of expression of the selected gene to induce an immune response. Where protective immunity is provided, the recombinant adenoviruses are considered to be vaccine compositions useful in preventing infection and/or recurrent disease.

Alternatively, or in addition, the vectors of the invention may contain a transgene encoding a peptide, polypeptide or protein which induces an immune response to a selected immunogen. The recombinant adenoviruses of this invention are expected to be highly efficacious at inducing cytolytic T cells and antibodies to the inserted heterologous antigenic protein expressed by the vector.

For example, immunogens may be selected from a variety of viral families. Example of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and human enteroviruses such as hepatitis A virus; and the genera aphthoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Another viral family includes the calcivirus family, which encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, RossRiver virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus, including Rubella virus. The flaviviridae family includes dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinin encephalomyelitis virus (pig), feline infectious peritonitis virus (cats), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis. In addition, the human coronaviruses include the putative causative agent of sudden acute respiratory syndrome (SARS). Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called HE or hemagglutin-esterose) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the rhabdovirus family, which includes the genera vesiculovirus (e.g., Vesicular

Stomatitis Virus), and the general lyssavirus (e.g., rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein. The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus, may be a suitable source of antigens. The paramyxovirus family

5 includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus), parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the

10 family orthomyxovirus and is a suitable source of antigen (e.g., the HA protein, the NI protein). The bunyavirus family includes the genera bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (puremala is a hemahagin fever virus), nairovirus (Nairobi sheep disease) and various unassigned bungaviruses. The arenavirus family provides a source of antigens against LCM and

15 Lassa fever virus. The reovirus family includes the genera reovirus, rotavirus (which causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever), Lebombo (humans), equine encephalosis, blue tongue.

The retrovirus family includes the sub-family oncorivirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLV I and HTLVII, lentivirinal (which includes human immunodeficiency virus (HIV), simian immunodeficiency virus (SIV), feline immunodeficiency virus (FIV), equine infectious anemia virus, and spumavirinal). Among the lentiviruses, many suitable antigens have been described and can readily be selected. Examples of suitable HIV and SIV antigens include, without limitation the gag, pol, Vif, Vpx, VPR, Env, Tat, Nef, and Rev proteins, as well as various fragments thereof. For example, suitable fragments of the Env protein may include any of its subunits such as the gp120, gp160, gp41, or smaller fragments thereof, e.g., of at least about 8 amino acids in length. Similarly, fragments of the tat protein may be selected. [See, US Patent 5,891,994 and US Patent 6,193,981.] See, also, the HIV and SIV proteins described in D.H. Barouch *et al*, *J. Virol.*, 75(5):2462-2467 (March 2001), and R.R. Amara, *et al*, *Science*, 292:69-74 (6 April 2001). In another example, the HIV and/or SIV immunogenic proteins or peptides may be used to form fusion proteins or other immunogenic molecules. See, e.g., the HIV-1 Tat and/or Nef fusion proteins and

immunization regimens described in WO 01/54719, published August 2, 2001, and WO 99/16884, published April 8, 1999. The invention is not limited to the HIV and/or SIV immunogenic proteins or peptides described herein. In addition, a variety of modifications to these proteins have been described or could readily be made by one of skill in the art. See, e.g., the modified gag protein that is described in US Patent 5,972,596. Further, any desired HIV and/or SIV immunogens may be delivered alone or in combination. Such combinations may include expression from a single vector or from multiple vectors. Optionally, another combination may involve delivery of one or more expressed immunogens with delivery of one or more of the immunogens in protein form. Such combinations are discussed in more detail below.

The papovavirus family includes the sub-family polyomaviruses (BKU and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis. The parvovirus includes family feline parvovirus (feline enteritis), feline panleucopeniavirus, canine parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family alphaherpesvirinae, which encompasses the genera simplexvirus (HSV_I, HSV_{II}), varicellovirus (pseudorabies, varicella zoster) and the sub-family betaherpesvirinae, which includes the genera cytomegalovirus (HCMV, muromegalovirus) and the sub-family gammaherpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitts lymphoma), infectious rhinotracheitis, Marek's disease virus, and rhabdovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus family includes equine arteritis virus and various Encephalitis viruses.

The viruses of the present invention may also carry immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell.

Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci; and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; pseudomonas, acinetobacteria and eikenella; melioidosis; salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (which causes chancroid); brucella; *Francisella tularensis* (which causes tularemia); yersinia (pasteurella); streptobacillus moniliformis and spirillum; Gram-positive bacilli include listeria monocytogenes; erysipelothrix rhusiopathiae; *Corynebacterium diphtheriae* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; treponematoses: yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include actinomycosis; nocardiosis; cryptococcosis, blastomycosis, histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis; sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever, Rocky Mountain spotted fever, Q fever, and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections.

Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis; toxoplasmosis; *Pneumocystis carinii*; Trichans; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes or flukes; and cestode (tapeworm) infections.

Many of these organisms and/or toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax), *Clostridium botulinum* and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fevers [filoviruses (e.g., Ebola, Marburg], and arenaviruses [e.g., Lassa, Machupo]), all of which are currently classified as Category A agents; *Coxiella burnetti* (Q fever);

Brucella species (brucellosis), *Burkholderia mallei* (glanders), *Burkholderia pseudomallei* (meloidosis), *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), *Staphylococcus* species and their toxins (enterotoxin B), *Chlamydia psittaci* (psittacosis), water safety threats (e.g., *Vibrio cholerae*, *Cryptosporidium parvum*), Typhus fever (*Rickettsia powazekii*), and viral encephalitis (alphaviruses, e.g., Venezuelan equine encephalitis; eastern equine encephalitis; western equine encephalitis); all of which are currently classified as Category B agents; and Nipah virus and hantaviruses, which are currently classified as Category C agents.

5 In addition, other organisms, which are so classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described herein are useful to deliver antigens from these organisms, viruses, their toxins or other by-products, which will prevent and/or treat infection or other adverse reactions with these biological agents.

10

15 Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In RA, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and Va-17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In MS, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-7 and Va-10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in MS. In scleroderma, several 20 specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and Va-16, Va-3C, Va-7, Va-14, Va-15, Va-16, Va-28 and Va-12. Thus, delivery of a chimeric adenovirus that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

25

30 C. Ad-Mediated Delivery Methods

The therapeutic levels, or levels of immunity, of the selected gene can be monitored to determine the need, if any, for boosters. Following an assessment of CD8+ T cell response, or optionally, antibody titers, in the serum, optional booster

immunizations may be desired. Optionally, the adenoviral vectors of the invention may be delivered in a single administration or in various combination regimens, e.g., in combination with a regimen or course of treatment involving other active ingredients or in a prime-boost regimen. A variety of such regimens have been

5 described in the art and may be readily selected.

For example, prime-boost regimens may involve the administration of a DNA (e.g., plasmid) based vector to prime the immune system to a second or further, booster, administration with a traditional antigen, such as a protein or a recombinant virus carrying the sequences encoding such an antigen. See, e.g., WO

10 00/11140, published March 2, 2000, incorporated by reference. Alternatively, an immunization regimen may involve the administration of a chimeric adenoviral vector of the invention to boost the immune response to a vector (either viral or DNA-based) carrying an antigen, or a protein. In still another alternative, an immunization regimen involves administration of a protein followed by booster with a vector

15 encoding the antigen.

In one embodiment, the invention provides a method of priming and boosting an immune response to a selected antigen by delivering a plasmid DNA vector carrying said antigen, followed by boosting with an adenoviral vector of the invention. In one embodiment, the prime-boost regimen involves the expression of

20 multiproteins from the prime and/or the boost vehicle. See, e.g., R.R. Amara, *Science*, 292:69-74 (6 April 2001) which describes a multiprotein regimen for expression of protein subunits useful for generating an immune response against HIV and SIV. For example, a DNA prime may deliver the Gag, Pol, Vif, VPX and Vpr and Env, Tat, and Rev from a single transcript. Alternatively, the SIV Gag, Pol and

25 HIV-1 Env is delivered in a recombinant adenovirus construct of the invention. Still other regimens are described in WO 99/16884 and WO 01/54719.

However, the prime-boost regimens are not limited to immunization for HIV or to delivery of these antigens. For example, priming may involve delivering with a first vector of the invention followed by boosting with a second

30 vector, or with a composition containing the antigen itself in protein form. In one example, the prime-boost regimen can provide a protective immune response to the virus, bacteria or other organism from which the antigen is derived. In another desired embodiment, the prime-boost regimen provides a therapeutic effect that can

be measured using convention assays for detection of the presence of the condition for which therapy is being administered.

The priming composition may be administered at various sites in the body in a dose dependent manner, which depends on the antigen to which the desired immune response is being targeted. The invention is not limited to the amount or

situs of injection(s) or to the pharmaceutical carrier. Rather, the regimen may

involve a priming and/or boosting step, each of which may include a single dose or

dosage that is administered hourly, daily, weekly or monthly, or yearly. As an

example, the mammals may receive one or two doses containing between about 10 µg

10 to about 50 µg of plasmid in carrier. A desirable amount of a DNA composition ranges between about 1 µg to about 10,000 µg of the DNA vector. Dosages may vary from about 1 µg to 1000 µg DNA per kg of subject body weight. The amount or site of delivery is desirably selected based upon the identity and condition of the mammal.

The dosage unit of the vector suitable for delivery of the antigen to the mammal is

15 described herein. The vector is prepared for administration by being suspended or dissolved in a pharmaceutically or physiologically acceptable carrier such as isotonic saline; isotonic salts solution or other formulations that will be apparent to those skilled in such administration. The appropriate carrier will be evident to those skilled in the art and will depend in large part upon the route of administration. The

20 compositions of the invention may be administered to a mammal according to the routes described above, in a sustained release formulation using a biodegradable biocompatible polymer, or by on-site delivery using micelles, gels and liposomes.

Optionally, the priming step of this invention also includes administering with the priming composition, a suitable amount of an adjuvant, such as are defined herein.

25 Preferably, a boosting composition is administered about 2 to about 27 weeks after administering the priming composition to the mammalian subject. The administration of the boosting composition is accomplished using an effective amount of a boosting composition containing or capable of delivering the same antigen as administered by the priming DNA vaccine. The boosting composition may be

30 composed of a recombinant viral vector derived from the same viral source (e.g., adenoviral sequences of the invention) or from another source. Alternatively, the “boosting composition” can be a composition containing the same antigen as encoded in the priming DNA vaccine, but in the form of a protein or peptide, which

composition induces an immune response in the host. In another embodiment, the boosting composition contains a DNA sequence encoding the antigen under the control of a regulatory sequence directing its expression in a mammalian cell, e.g., vectors such as well-known bacterial or viral vectors. The primary requirements of the boosting composition are that the antigen of the composition is the same antigen, or a cross-reactive antigen, as that encoded by the priming composition.

In another embodiment, the adenoviral vectors of the invention are also well suited for use in a variety of other immunization and therapeutic regimens.

Such regimens may involve delivery of adenoviral vectors of the invention

10 simultaneously or sequentially with Ad vectors of different serotype capsids, regimens in which adenoviral vectors of the invention are delivered simultaneously or sequentially with non-Ad vectors, regimens in which the adenoviral vectors of the invention are delivered simultaneously or sequentially with proteins, peptides, and/or other biologically useful therapeutic or immunogenic compounds. Such uses will be
15 readily apparent to one of skill in the art.

IV. Simian Adenovirus 18 Sequences

The invention provides nucleic acid sequences and amino acid sequences of Ad SA18, which are isolated from the other viral material with which they are

20 associated in nature. These sequences are useful in preparing heterologous molecules containing the nucleic acid sequences and amino acid sequences, and regions or fragments thereof as are described herein, viral vectors which are useful for a variety of purposes, including the constructs and compositions, and such methods as are described herein for the chimeric adenoviruses, including, e.g., in host cells for
25 production of viruses requiring adenoviral helper functions, as delivery vehicles for heterologous molecules such as those described herein. These sequences are also useful in generating the chimeric adenoviruses of the invention.

A. Nucleic Acid Sequences

The SA18 nucleic acid sequences of the invention include nucleotides

30 SEQ ID NO: 12, nt 1 to 31967. See, Sequence Listing, which is incorporated by reference herein. The nucleic acid sequences of the invention further encompass the strand which is complementary to the sequences of SEQ ID NO: 12, as well as the RNA and cDNA sequences corresponding to the sequences of these sequences figures

and their complementary strands. Further included in this invention are nucleic acid sequences which are greater than 95 to 98%, and more preferably about 99 to 99.9% homologous or identical to the Sequence Listing. Also included in the nucleic acid sequences of the invention are natural variants and engineered modifications of the

5 sequences provided in SEQ ID NO: 12 and their complementary strands. Such modifications include, for example, labels that are known in the art, methylation, and substitution of one or more of the naturally occurring nucleotides with a degenerate nucleotide.

The invention further encompasses fragments of the sequences of
10 SA18, their complementary strand, cDNA and RNA complementary thereto. Suitable fragments are at least 15 nucleotides in length, and encompass functional fragments, i.e., fragments which are of biological interest. For example, a functional fragment can express a desired adenoviral product or may be useful in production of recombinant viral vectors. Such fragments include the gene sequences and fragments
15 listed in the tables below.

The following tables provide the transcript regions and open reading frames in the simian adenovirus sequences of the invention. For certain genes, the transcripts and open reading frames (ORFs) are located on the strand complementary to that presented in SEQ ID NO: 12. See, e.g., E2b, E4 and E2a. The calculated
20 molecular weights of the encoded proteins are also shown.

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
		start	End	M.W.
ITR		1	180	
E1a	13S	916	1765	27264
	12S	916	1765	24081
E1b	Small T	1874	2380	19423
	LargeT	2179	3609	52741
	IX	3678	4079	13701
E2b	IVa2	5478	4126	51295
	Polymerase	13745	5229	128392
	PTP	13745	8597	75358
	Agnoprotein	8007	8705	23610
L1	52/55 kD	10788	11945	43416
	IIIa	11966	13699	63999
L2	Penton	13796	15322	57166
	VII	15328	15873	20352
	V	15920	17050	42020
L3	VI	17348	18154	29222
	Hexon	18257	21010	102912
	Endoprotease	21029	21640	23015

Adenovirus Gene Region	Protein	Ad SA18, SEQ ID NO:12		
2a	DBP	23147	21711	53626
L4	100kD	23175	25541	87538
	22 kD homolog	25204	25797	22206
	33 kD homolog	25204	26025	24263
	VIII	26107	26817	25490
E3	Orf #1	26817	27125	11814
L5	Fiber	27192	29015	65455
E4	Orf 6/7	30169	29067	13768
	Orf 6	30169	29303	33832
	Orf 4	30464	30099	14154
	Orf 3	30816	30466	13493
	Orf 2	31205	30813	14698
	Orf 1	31608	31231	14054
ITR		31788	31967	

The SA18 adenoviral nucleic acid sequences are useful as therapeutic and immunogenic agents and in construction of a variety of vector systems and host cells. Such vectors are useful for any of the purposes described above for the chimeric adenovirus. Additionally, these SA18 sequences and products may be used alone or in combination with other adenoviral sequences or fragments, or in combination with elements from other adenoviral or non-adenoviral sequences. The adenoviral sequences of the invention are also useful as antisense delivery vectors, gene therapy vectors, or vaccine vectors, and in methods of using same. Thus, the invention further provides nucleic acid molecules, gene delivery vectors, and host cells which contain the Ad sequences of the invention.

For example, the invention encompasses a nucleic acid molecule containing simian Ad ITR sequences of the invention. In another example, the invention provides a nucleic acid molecule containing simian Ad sequences of the

invention encoding a desired Ad gene product. Still other nucleic acid molecule constructed using the sequences of the invention will be readily apparent to one of skill in the art, in view of the information provided herein.

In one embodiment, the simian Ad gene regions identified herein may 5 be used in a variety of vectors for delivery of a heterologous molecule to a cell. Examples of such molecules and methods of delivery are provided in Section III herein. For example, vectors are generated for expression of an adenoviral capsid protein (or fragment thereof) for purposes of generating a viral vector in a packaging host cell. Such vectors may be designed for expression in trans. Alternatively, such 10 vectors are designed to provide cells which stably contain sequences which express desired adenoviral functions, e.g., one or more of E1a, E1b, the terminal repeat sequences, E2a, E2b, E4, E4ORF6 region.

In addition, the adenoviral gene sequences and fragments thereof are useful for providing the helper functions necessary for production of helper- 15 dependent viruses (e.g., adenoviral vectors deleted of essential functions or adeno-associated viruses (AAV)). For such production methods, the simian adenoviral sequences of the invention are utilized in such a method in a manner similar to those described for the human Ad. However, due to the differences in sequences between the simian adenoviral sequences of the invention and those of human Ad, the use of 20 the sequences of the invention essentially eliminate the possibility of homologous recombination with helper functions in a host cell carrying human Ad E1 functions, e.g., 293 cells, which may produce infectious adenoviral contaminants during rAAV production.

Methods of producing rAAV using adenoviral helper functions have 25 been described at length in the literature with human adenoviral serotypes. See, e.g., US Patent 6,258,595 and the references cited therein. See, also, US Patent 5,871,982; WO 99/14354; WO 99/15685; WO 99/47691. These methods may also be used in production of non-human serotype AAV, including non-human primate AAV 30 serotypes. The simian adenoviral gene sequences of the invention which provide the necessary helper functions (e.g., E1a, E1b, E2a and/or E4 ORF6) can be particularly useful in providing the necessary adenoviral function while minimizing or eliminating the possibility of recombination with any other adenoviruses present in the rAAV- packaging cell which are typically of human origin. Thus, selected genes or open

reading frames of the adenoviral sequences of the invention may be utilized in these rAAV production methods.

Alternatively, recombinant adenoviral simian vectors of the invention may be utilized in these methods. Such recombinant adenoviral simian vectors may

5 include, e.g., a hybrid simian Ad/AAV in which simian Ad sequences flank a rAAV expression cassette composed of, e.g., AAV 3' and/or 5' ITRs and a transgene under the control of regulatory sequences which control its expression. One of skill in the art will recognize that still other simian adenoviral vectors and/or gene sequences of the invention will be useful for production of rAAV and other viruses dependent upon

10 adenoviral helper.

In still another embodiment, nucleic acid molecules are designed for delivery and expression of selected adenoviral gene products in a host cell to achieve a desired physiologic effect. For example, a nucleic acid molecule containing sequences encoding an adenovirus Ela protein of the invention may be delivered to a

15 subject for use as a cancer therapeutic. Optionally, such a molecule is formulated in a lipid-based carrier and preferentially targets cancer cells. Such a formulation may be combined with other cancer therapeutics (e.g., cisplatin, taxol, or the like). Still other uses for the adenoviral sequences provided herein will be readily apparent to one of skill in the art.

20 In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and non-viral vector systems for in vitro, ex vivo or in vivo delivery of therapeutic and immunogenic molecules, including any of those identified as being deliverable via the chimeric adenoviruses of the invention. For example, the simian Ad genome of the

25 invention can be utilized in a variety of rAd and non-rAd vector systems. Such vectors systems may include, e.g., plasmids, lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adeno-associated viral systems, among others. Selection of these vector systems is not a limitation of the present invention.

The invention further provides molecules useful for production of the

30 simian and simian-derived proteins of the invention. Such molecules which carry polynucleotides including the simian Ad DNA sequences of the invention can be in the form of a vector.

B. Simian Adenoviral Proteins of the Invention

The invention further provides gene products of the above adenoviruses, such as proteins, enzymes, and fragments thereof, which are encoded by the adenoviral nucleic acids of the invention. The invention further encompasses 5 SA18 proteins, enzymes, and fragments thereof, having the amino acid sequences encoded by these nucleic acid sequences which are generated by other methods. Such proteins include those encoded by the open reading frames identified in the tables above, and fragments thereof.

Thus, in one aspect, the invention provides unique simian adenoviral 10 proteins which are substantially pure, *i.e.*, are free of other viral and proteinaceous proteins. Preferably, these proteins are at least 10% homogeneous, more preferably 60% homogeneous, and most preferably 95% homogeneous.

In one embodiment, the invention provides unique simian-derived capsid proteins. As used herein, a simian-derived capsid protein includes any 15 adenoviral capsid protein that contains a SA18 capsid protein or a fragment thereof, as defined above, including, without limitation, chimeric capsid proteins, fusion proteins, artificial capsid proteins, synthetic capsid proteins, and recombinantly capsid proteins, without limitation to means of generating these proteins.

Suitably, these simian-derived capsid proteins contain one or more 20 SA18 regions or fragments thereof (e.g., a hexon, penton, fiber or fragment thereof) in combination with capsid regions or fragments thereof of different adenoviral serotypes, or modified simian capsid proteins or fragments, as described herein. A "modification of a capsid protein associated with altered tropism" as used herein includes an altered capsid protein, *i.e.*, a penton, hexon or fiber protein region, or 25 fragment thereof, such as the knob domain of the fiber region, or a polynucleotide encoding same, such that specificity is altered. The simian-derived capsid may be constructed with one or more of the simian Ad of the invention or another Ad serotypes which may be of human or non-human origin. Such Ad may be obtained from a variety of sources including the ATCC, commercial and academic sources, or 30 the sequences of the Ad may be obtained from GenBank or other suitable sources.

The amino acid sequences of the simian adenoviruses penton proteins of the invention are provided herein. The AdSA18 penton protein is provided in SEQ ID NO: 13. Suitably, any of these penton proteins, or unique fragments thereof, may

be utilized for a variety of purposes. Examples of suitable fragments include the penton having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. Further, the penton protein may be modified for a variety of purposes known to those of skill in the art.

The invention further provides the amino acid sequences of the hexon protein of SA18, SEQ ID NO:14. Suitably, this hexon protein, or unique fragments thereof, may be utilized for a variety of purposes. Examples of suitable fragments include the hexon having N-terminal and/or C-terminal truncations of about 50, 100, 150, 200, 300, 400, or 500 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 14. Other suitable fragments include shorter internal, C-terminal, or N-terminal fragments. For example, one suitable fragment the loop region (domain) of the hexon protein, designated DE1 and FG1, or a hypervariable region thereof. Such fragments include the regions spanning amino acid residues about 125 to 443; about 138 to 441, or smaller fragments, such as those spanning about residue 138 to residue 163; about 170 to about 176; about 195 to about 203; about 233 to about 246; about 253 to about 264; about 287 to about 297; about 404 to about 430, about 430 to 550, about 545 to 650; of the simian hexon proteins, with reference to SEQ ID NO: 14. Other suitable fragments may be readily identified by one of skill in the art. Further, the hexon protein may be modified for a variety of purposes known to those of skill in the art. Because the hexon protein is the determinant for serotype of an adenovirus, such artificial hexon proteins would result in adenoviruses having artificial serotypes. Other artificial capsid proteins can also be constructed using the chimp Ad penton sequences and/or fiber sequences of the invention and/or fragments thereof.

In one example, it may be desirable to generate an adenovirus having an altered hexon protein utilizing the sequences of a hexon protein of the invention. One suitable method for altering hexon proteins is described in US Patent 5,922,315, which is incorporated by reference. In this method, at least one loop region of the adenovirus hexon is changed with at least one loop region of another adenovirus serotype. Thus, at least one loop region of such an altered adenovirus hexon protein is a simian Ad hexon loop region of the invention. In one embodiment, a loop region

of the SA18 hexon protein is replaced by a loop region from another adenovirus serotype. In another embodiment, the loop region of the SA18 hexon is used to replace a loop region from another adenovirus serotype. Suitable adenovirus serotypes may be readily selected from among human and non-human serotypes, as described herein. SA18 is selected for purposes of illustration only; the other simian Ad hexon proteins of the invention may be similarly altered, or used to alter another Ad hexon. The selection of a suitable serotype is not a limitation of the present invention. Still other uses for the hexon protein sequences of the invention will be readily apparent to those of skill in the art.

The invention further encompasses the fiber proteins of the simian adenoviruses of the invention. The fiber protein of AdSA18 has the amino acid sequence of SEQ ID NO: 15. Suitably, this fiber protein, or unique fragments thereof, may be utilized for a variety of purposes. One suitable fragment is the fiber knob, which spans about amino acids 247 to 425 of SEQ ID NO: 15. Examples of other suitable fragments include the fiber having N-terminal and/or C-terminal truncations of about 50, 100, 150, or 200 amino acids, based upon the amino acid numbering provided above and in SEQ ID NO: 15. Still other suitable fragments include internal fragments. Further, the fiber protein may be modified using a variety of techniques known to those of skill in the art.

The invention further encompasses unique fragments of the proteins of the invention which are at least 8 amino acids in length. However, fragments of other desired lengths can be readily utilized. In addition, the invention encompasses such modifications as may be introduced to enhance yield and/or expression of an SA18 gene product, e.g., construction of a fusion molecule in which all or a fragment of the SA18 gene product is fused (either directly or via a linker) with a fusion partner to enhance. Other suitable modifications include, without limitation, truncation of a coding region (e.g., a protein or enzyme) to eliminate a pre- or pro-protein ordinarily cleaved and to provide the mature protein or enzyme and/or mutation of a coding region to provide a secretable gene product. Still other modifications will be readily apparent to one of skill in the art. The invention further encompasses proteins having at least about 95% to 99% identity to the SA18 proteins provided herein.

As described herein, vectors of the invention containing the adenoviral capsid proteins of the invention are particularly well suited for use in applications in which the neutralizing antibodies diminish the effectiveness of other Ad serotype based vectors, as well as other viral vectors. The rAd vectors of the invention are particularly advantageous in readministration for repeat gene therapy or for boosting immune response (vaccine titers). Examples of such regimens are provided herein.

Under certain circumstances, it may be desirable to use one or more of the SA18 gene products (e.g., a capsid protein or a fragment thereof) to generate an antibody. The term "an antibody," as used herein, refers to an immunoglobulin molecule which is able to specifically bind to an epitope. Thus, the antibodies of the invention bind, preferably specifically and without cross-reactivity, to a SA18 epitope. The antibodies in the present invention exist in a variety of forms including, for example, high affinity polyclonal antibodies, monoclonal antibodies, synthetic antibodies, chimeric antibodies, recombinant antibodies and humanized antibodies. Such antibodies originate from immunoglobulin classes IgG, IgM, IgA, IgD and IgE.

Such antibodies may be generated using any of a number of methods known in the art. Suitable antibodies may be generated by well-known conventional techniques, e.g. Kohler and Milstein and the many known modifications thereof. Similarly desirable high titer antibodies are generated by applying known recombinant techniques to the monoclonal or polyclonal antibodies developed to these antigens [see, e.g., PCT Patent Application No. PCT/GB85/00392; British Patent Application Publication No. GB2188638A; Amit *et al.*, 1986 *Science*, 233:747-753; Queen *et al.*, 1989 *Proc. Nat'l. Acad. Sci. USA*, 86:10029-10033; PCT Patent Application No. PCT/WO9007861; and Riechmann *et al.*, *Nature*, 332:323-327 (1988); Huse *et al.*, 1988a *Science*, 246:1275-1281]. Alternatively, antibodies can be produced by manipulating the complementarity determining regions of animal or human antibodies to the antigen of this invention. See, e.g., E. Mark and Padlin, "Humanization of Monoclonal Antibodies", Chapter 4, *The Handbook of Experimental Pharmacology*, Vol. 113, *The Pharmacology of Monoclonal Antibodies*, Springer-Verlag (June, 1994); Harlow *et al.*, 1999, *Using Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, NY; Harlow *et al.*, 1989, *Antibodies: A Laboratory Manual*, Cold Spring Harbor, New York; Houston *et al.*, 1988, *Proc.*

Natl. Acad. Sci. USA 85:5879-5883; and Bird *et al.*, 1988, *Science* 242:423-426. Further provided by the present invention are anti-idiotype antibodies (Ab2) and anti-anti-idiotype antibodies (Ab3). See, e.g., M. Wettendorff *et al.*, "Modulation of anti-tumor immunity by anti-idiotypic antibodies." In *Idiotypic Network and Diseases*, ed. by J. Cerny and J. Hiernaux, 1990 *J. Am. Soc. Microbiol.*, Washington DC: pp. 203-229]. These anti-idiotype and anti-anti-idiotype antibodies are produced using techniques well known to those of skill in the art. These antibodies may be used for a variety of purposes, including diagnostic and clinical methods and kits.

Under certain circumstances, it may be desirable to introduce a detectable label or a tag onto a SA18 gene product, antibody or other construct of the invention. As used herein, a detectable label is a molecule which is capable, alone or upon interaction with another molecule, of providing a detectable signal. Most desirably, the label is detectable visually, e.g. by fluorescence, for ready use in immunohistochemical analyses or immunofluorescent microscopy. For example, suitable labels include fluorescein isothiocyanate (FITC), phycoerythrin (PE), allophycocyanin (APC), coriphosphine-O (CPO) or tandem dyes, PE-cyanin-5 (PC5), and PE-Texas Red (ECD). All of these fluorescent dyes are commercially available, and their uses known to the art. Other useful labels include a colloidal gold label. Still other useful labels include radioactive compounds or elements. Additionally, labels include a variety of enzyme systems that operate to reveal a colorimetric signal in an assay, e.g., glucose oxidase (which uses glucose as a substrate) releases peroxide as a product which in the presence of peroxidase and a hydrogen donor such as tetramethyl benzidine (TMB) produces an oxidized TMB that is seen as a blue color. Other examples include horseradish peroxidase (HRP) or alkaline phosphatase (AP), and hexokinase in conjunction with glucose-6-phosphate dehydrogenase which reacts with ATP, glucose, and NAD⁺ to yield, among other products, NADH that is detected as increased absorbance at 340 nm wavelength.

Other label systems that are utilized in the methods of this invention are detectable by other means, e.g., colored latex microparticles [Bangs Laboratories, Indiana] in which a dye is embedded are used in place of enzymes to form conjugates with the target sequences provide a visual signal indicative of the presence of the resulting complex in applicable assays.

Methods for coupling or associating the label with a desired molecule are similarly conventional and known to those of skill in the art. Known methods of label attachment are described [see, for example, Handbook of Fluorescent probes and Research Chemicals, 6th Ed., R. P. M. Haugland, Molecular Probes, Inc.,

5 Eugene, OR, 1996; Pierce Catalog and Handbook, Life Science and Analytical Research Products, Pierce Chemical Company, Rockford, IL, 1994/1995]. Thus, selection of the label and coupling methods do not limit this invention.

The sequences, proteins, and fragments of the invention may be produced by any suitable means, including recombinant production, chemical synthesis, or other synthetic means. Suitable production techniques are well known to those of skill in the art. See, e.g., Sambrook et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternatively, peptides can also be synthesized by the well known solid phase peptide synthesis methods (Merrifield, *J. Am. Chem. Soc.*, 85:2149 (1962); Stewart and 10 Young, Solid Phase Peptide Synthesis (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

In addition, one of skill in the art will readily understand that the Ad sequences of the invention can be readily adapted for use for a variety of viral and 20 non-viral vector systems for *in vitro*, *ex vivo* or *in vivo* delivery of therapeutic and immunogenic molecules. For example, in one embodiment, the simian Ad capsid proteins and other simian adenovirus proteins described herein are used for non-viral, protein-based delivery of genes, proteins, and other desirable diagnostic, therapeutic and immunogenic molecules. In one such embodiment, a protein of the invention is 25 linked, directly or indirectly, to a molecule for targeting to cells with a receptor for adenoviruses. Preferably, a capsid protein such as a hexon, penton, fiber or a fragment thereof having a ligand for a cell surface receptor is selected for such targeting. Suitable molecules for delivery are selected from among the therapeutic molecules described herein and their gene products. A variety of linkers including, lipids, 30 polyLys, and the like may be utilized as linkers. For example, the simian penton protein may be readily utilized for such a purpose by production of a fusion protein using the simian penton sequences in a manner analogous to that described in Medina-Kauwe LK, et al, *Gene Ther.* 2001 May; 8(10):795-803 and Medina-Kauwe

LK, et al, *Gene Ther.* 2001 Dec; 8(23): 1753-1761. Alternatively, the amino acid sequences of simian Ad protein IX may be utilized for targeting vectors to a cell surface receptor, as described in US Patent Appln 20010047081. Suitable ligands include a CD40 antigen, an RGD-containing or polylysine-containing sequence, and the like. Still other simian Ad proteins, including, e.g., the hexon protein and/or the fiber protein, may be used for these and similar purposes.

5 Still other adenoviral proteins of the invention may be used as alone, or in combination with other adenoviral protein, for a variety of purposes which will be readily apparent to one of skill in the art. In addition, still other uses for the
10 adenoviral proteins of the invention will be readily apparent to one of skill in the art.

10 The compositions of this invention include vectors that deliver a heterologous molecule to cells, either for therapeutic or vaccine purposes. Such vectors, containing simian adenovirus DNA of SA18 and a minigene, can be constructed using techniques such as those described herein for the chimeric
15 adenoviruses and such techniques as are known in the art. Alternatively, SA19 may be a source for sequences of the chimeric adenoviruses are described herein.

20 The following examples illustrate construction and use of several chimeric viruses, including Pan5/C1, hu5/Pan7 and hu5/SV25, and Pan6/Pan7.
20 However, these chimera are illustrative only and are not intended to limit the invention to those illustrated embodiments.

Example 1 - Construction of Pan5/C1 Chimeric Simian Viruses

25 Five different adenoviruses initially isolated from the chimpanzee, AdC68 [US Patent 6,083,716], AdPan5, AdPan7, AdPan6 and AdC1 [US Patent 6,083,716] have been sequenced. See, International Application No. PCT/US02/33645, filed November 2002 for the sequences of Pan5 [SEQ ID NO:1], Pan7 [SEQ ID NO:3], and Pan6 [SEQ ID NO:2]. This application also provides sequences for SV1, SV25 and SV39 [SEQ ID No. 4, 5, 6, respectively]. Sequence comparison of the capsid
30 protein sequences predicted that AdC1 clearly belonged to a different serological subgroup than the other four chimpanzee derived adenoviruses.

However, attempts to cultivate AdC1 in HEK293 cells revealed it to be fastidious in its growth characteristics (data not shown) and therefore possibly

unsuitable for use as a vector using the currently available E1 complementing cell lines. However, because of the obvious sequence dissimilarity of AdC1 capsid protein sequence from the other chimpanzee derived adenoviruses (as well as the huAd5), chimeric adenovirus vectors were generated with the capsid characteristics of AdC1. In view of the above-mentioned drawbacks associated with only making hexon changes, more extensive replacements were made in the chimera described herein, i.e., construction of chimeras where the replacement went beyond just the hexon, to achieve two goals. The first was to determine whether making extended replacements would allow for the rescue of viruses containing hexons of unrelated serotypes that may not otherwise be amenable to rescue. The second goal was to test whether the growth characteristics of adenovirus vectors such as AdPan5, that have been found in our laboratory to be able to be grown to high titer for the purpose of manufacture, would also be present in the chimeric virus, particularly when the hexon (and other capsid proteins) are derived from a virus such as AdC1 that are difficult to grow to a high yield in cell lines such as HEK293. An added bonus of extending the replacement to include the fiber protein would be to further increase the antigenic dissimilarity to beyond that afforded by a hexon change alone.

As an alternative to obtaining purified virus as source for adenoviral DNA to sequence, we have resorted to cloning restriction fragments of viral DNA obtained from infected cells ("Hirt prep"). The first adenovirus we have sequenced in this way is Simian Adenovirus. EcoRI digestion of the Simian Adenovirus yielded 7 fragments. Shotgun cloning yielded clones of the 5 internal fragments, which were cloned and sequenced. Completion of the sequencing was carried out by walking towards each of the ends of the genome. The map of the genome is shown in Figure 1.

A. *Construction of Two Pan5/C1 Chimeric Plasmids*

The overall approach towards constructing chimeric viruses was to first assemble the complete E1 deleted virus DNA into a single plasmid flanked by recognition sites for the restriction enzyme SwaI, digest the plasmid DNA with SwaI to release the virus DNA ends, and transfet the DNA into HEK293 cells to determine whether viable chimeric adenovirus could be rescued. Two chimeric virus plasmids were constructed, p5C1short and p5C1long.

The plasmid p5C1short harbors an E1 deleted Pan5 virus where an internal 15226 bp segment (18332 – 33557) has been replaced by a functionally analogous 14127 bp (18531 – 32657) from AdC1. This results in the replacement of the Pan5 proteins hexon, endoprotease, DNA binding protein, 100 kD scaffolding protein, 33 kD protein, protein VIII, and fiber, as well as the entire E3 region, with the homologous segment from AdC1. The Clal site at the left end of the AdC1 fragment is at the beginning of the hexon gene and the resulting protein is identical to the C1 hexon. The EcoRI site which constitutes the right end of the AdC1 fragment is within the E4 orf 7 part of the AdC1. The right end was ligated to a PCR generated right end fragment from AdPan5 such that the regenerated orf 7-translation product is chimeric between AdPan5 and AdC1.

The plasmid p5C1long harbors an E1 deleted Pan5 virus where an internal 25603 bp segment (7955 – 33557) has been replaced by a functionally analogous 24712 bp (7946 – 32657) from AdC1. This results in the replacement of the AdPan5 pre-terminal protein, 52/55 kD protein, penton base protein, protein VII, Mu, and protein VI with those from AdC1 in addition to those replaced in p5C1short. The Ascl site at the left end of the AdC1 fragment is at the beginning of the DNA polymerase gene and results in a chimeric protein where the first 165 amino acids of the AdPan5 DNA polymerase has been replaced by a 167 amino acid segment from AdC1 DNA polymerase. In this N-terminal region, the homology between the AdPan5 and AdC1 DNA polymerase proteins is 81% (72% identity).

The plasmid pDVP5Mlu which contains the left end of AdPan5 was used as the starting plasmid for the chimeric vector construction.

The plasmid pDVP5Mlu was made as follows. A synthetic DNA fragment harboring recognition sites for the restriction enzymes SmaI, MluI, EcoRI and EcoRV respectively was ligated into pBR322 digested with EcoRI and NdeI so as to retain the origin of replication and the beta-lactamase gene. The left end of Pan5 extending to the MluI site (15135 bp) was cloned into this plasmid between the SmaI and MluI sites. The E1 gene was functionally deleted and replaced by a DNA fragment harboring recognition sites for the extremely rare cutter restriction enzyme sites I-CeuI and PI-SceI). The 2904 base pairs of the right end of Pan-5 was PCR amplified using the primers P5L [GCG CAC GCG TCT CTA TCG ATG AAT TCC ATT GGT GAT GGA CAT GC, SEQ ID NO:7] and PSITR [GCG CAT TTA AAT

CAT CAT CAA TAA TAT ACC TCA AAC, SEQ ID NO:8] using Tgo polymerase (Roche). The PCR product was cut with MluI and SwaI, and cloned between MluI and EcoRV of pDVPSMlu to yield pPan5Mlu+RE. A 3193 bp fragment extending from the MluI site (15135) to the ClaI (18328) site of Pan5 was then inserted between

5 the same sites of pPan5Mlu+RE to yield pPan5Cla+RE. The 3671 bp ClaI (18531) to EcoRI (22202) fragment of the adenovirus C1 was cloned into pPan5Cla+RE between ClaI (16111) and EcoRI (16116) to yield pPan5C1delRI. The 10452 bp internal EcoRI fragment of the adenovirus C1 (22202 – 32653) was cloned into the EcoRI site of pPan5C1delRI to yield p5C1short. To construct p5C1long, the AdC1 replacement
10 was further extended by replacing the AscI – ClaI 10379 bp fragment of AdPan5 in p5C1short with the AdC1 AscI – ClaI 10591 bp fragment. Finally a green fluorescent protein (GFP) expression cassette was inserted into both p5C1short and p5C1long between the I-CeuI and PI-SceI sites to yield p5C1shortGFP and p5C1longGFP respectively.

15 B. *Rescue of chimeric Pan5/C1 recombinant vector adenoviruses*

The plasmids p5C1shortGFP and p5C1longGFP were digested with the restriction enzyme SwaI and transfected into HEK 293 cells. A typical adenovirus induced cytopathic effect was observed. The rescue of recombinant chimeric adenovirus from the p5C1longGFP transfection was confirmed by collecting the
20 supernatant from the transfection and re-infecting fresh cells which were found to be transduced as determined by GFP expression. Viral DNA prepared from the chimeric recombinant virus was digested with several restriction enzymes and found to have the expected pattern on electrophoresis (data not shown).

The chimeric adenoviral construct with the shorter replacement
25 p5C1short encodes the C1 proteins hexon and fiber as well as the intervening open reading frames for endoprotease, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, and protein VIII. (The E3 region is also included within this region but is unlikely to impact on the viability of the chimeric virus). When the replacement was extended to include the additional AdC1 proteins pTP (pre-terminal
30 protein), 52/55 kDa protein, penton base, protein VII, Mu, and protein VI, there was no difficulty in rescuing viable chimeric virus. In this experiment, the chimeric adenovirus construction strategy utilized the presence of AscI and ClaI restriction

enzyme sites present on the genes for DNA polymerase and hexon respectively on both AdPan5 and AdC1.

The reasons for the relatively higher yield of the chimeric virus compared to the wild-type AdC1 virus are not clear. In the growth of the 5C1 chimeric virus in 293 cells, the adenoviral early region gene products of E1 and E4 are derived from Ad5 and AdPan5 respectively. The E1 and E4 gene products bind, regulate and de-repress several cellular transcription complexes and coordinate their activity towards viral multiplication. Thus it is possible that the E1 gene products supplied in *trans* from the 293 cells and the E4 gene products from AdPan5 are more optimal in the human 293 cell background than are the equivalent AdC1 gene products. This may also apply to the activity of the major late promoter whose activity is responsible for the transcription of the capsid protein genes. In the chimeric virus, the major late promoter, and the protein IVa2 which transactivates it, are derived from AdPan5. However the E2 gene products required for adenoviral DNA replication pTP and single-stranded DNA – binding protein are derived from AdC1. The adenoviral DNA polymerase, which complexes with pTP, is chimeric in Ad5C1 but mostly AdPan5 derived.

Two plasmids have been constructed where the structural proteins derive from the chimpanzee adenovirus Pan 7 and Simian adenovirus 25 respectively, and the flanking sequences are derived from human Ad5 (the commonly used vector strain). The Adhu5-Pan7 chimeric adenovirus has been rescued, demonstrating that the chimeric virus construction method used to derive the Pan 5 – C1 chimeric virus is broadly applicable. Fig. 2 provides the map of the recombinant Adhu5-SV25 chimeric virus. The portion of the genome replaced by DNA from Pan7 is indicated.

Example 2 – Pan5 - C1 chimeric vector of invention as a delivery vehicle for immunogenic compositions

A Pan 5 (Simian adenovirus 22, a subgroup E adenovirus, also termed C5) - C1 (Simian adenovirus 21, a subgroup B adenovirus) chimeric expressing the Ebola virus (Zaire) glycoprotein (C5C1-CMVGP) was constructed as a model antigen in order to test the efficacy of the vector C5C1-CMVGP as a vaccine; this vector has been compared it to the Adhu5 based vector (H5-CMVGP). Compared to H5-

CMVGP, the C5C1-CMVGP vector yielded only a slightly decreased level of GP expression in transduced A549 cells.

Thereafter, GP-specific T cell and B cell responses elicited in B10BR mice vaccinated with either H5-CMVGP or CSC1-CMVGP vectors were compared. The 5 C5C1-CMVGP vector appeared to induce lower frequencies of gamma interferon producing CD8+ T cells with kinetics slower than the H5-CMVGP vector as determined by intracellular cytokine staining using a H-2k restricted GP-specific peptide as stimulant. The total IgG response to GP, measured by ELISA, was equivalent in serum from mice vaccinated with the C5C1-CMVGP or the H5- 10 CMVGP vectors. However, the C5C1-CMVGP vector induced a more potent Th1 type response while the H5-CMVGP vector stimulated a more balanced Th1/Th2 type response.

Example 3 – Generation of Chimeric Pan6/Pan7 Vectors

15 A panel of GFP expressing vectors were generated. This panel includes vectors that are chimeric between Pan 6 and Pan 7 where (a) the hexon protein of Pan 7 was replaced by that of Pan 6 (termed C767), (b), the fiber protein of Pan7 was replaced by that of Pan 6 (termed C776), (c) both the hexon and fiber proteins of the Pan 7 vector have been replaced by those from Pan 6 (termed C766).

20 The chimeric virus termed C767 was constructed essentially as described above for the C5C1C5 virus in Example 1. However, due to substantial homology between the Pan6 and Pan7 sequences 5' to the hexon sequence, it was not necessary to substitute the 5' end of the genome between the penton and the pol gene.

25 The chimeric vector C767 was compared to the C776, C766, the parent C6, and the parent C7, each expressing GFP.

30 Balb/C mice (25 per group) were immunized intramuscularly with either Pan 6 or Pan 7 (10^{10} particles). Re-administration (10^{11} particles i.v., by tail vein injection) was attempted 3 weeks later using each of the five GFP expressing vectors (C6-GFP, C7-GFP, and the three chimeric vectors). Three days later the level of liver transduction was estimated qualitatively by examining liver sections for the presence of GFP expression and quantitatively by estimating copies of GFP DNA by Taqman analysis. Administration of either one of the two chimpanzee adenovirus vectors does not affect the transduction efficiency of the other vector, while re-administration of

the same vector is severly compromised. The data showed that antibodies to both hexon and fiber are important in preventing re-administration of adenoviral vectors.

5

All publications cited in this specification are incorporated herein by reference. While the invention has been described with reference to a particularly preferred embodiment, it will be appreciated that modifications can be made without departing from the spirit of the invention. Such modifications are intended to fall within the scope of the appended claims.

What is claimed is:

1. A method of efficiently culturing a chimeric adenovirus in a selected host cell, said chimeric adenovirus being from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the steps of:

(a) generating a chimeric adenovirus comprising:

(i) adenovirus sequences of the left terminal end and right terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the 5' inverted terminal repeat (ITRs), and said right end region comprising the 3' inverted terminal repeat (ITRs); and

(ii) the internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the late genes encoding the penton, hexon, and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', a left terminal region of the first adenovirus, the internal region of the parental adenovirus, and the right terminal region of the first adenovirus; and

b) culturing said chimeric adenovirus in the presence of functional adenovirus E1a, E1b, and E4 ORF6 genes from the first adenovirus or from an adenovirus serotype which transcomplements the first adenovirus, and further in the presence of necessary adenoviral structural genes from the left end of the adenovirus.

2. The method according to claim 1, wherein the internal region of the parental adenovirus further comprises one or more functional adenovirus genes selected from the group consisting of Endoprotease open reading frame, DNA binding protein, 100 kDa scaffolding protein, 33 kDa protein, protein VIII, pTP, 52/55 kDa protein, protein VII, Mu and protein VI.

3. The method according to claim 1, wherein the polymerase, terminal protein and 52/55 kDa protein functions are provided in *trans*.

4. The method according to claim 1, wherein the first adenovirus further comprises the polymerase, terminal protein and 52/55 kDa protein functions.
5. The method according to claim 1, wherein the chimeric adenovirus comprises the adenoviral late genes 1, 2, 3, 4, and 5 of the parental adenovirus.
6. The method according to claim 1, wherein the selected host cell stably contains one or more of the adenovirus E1a, E1b or E4 ORF6 functions.
7. The method according to claim 1, wherein the chimeric adenovirus comprises one or more of the adenovirus E1a, E1b or E4 ORF6 of the first adenovirus.
8. The method according to claim 1, wherein the first adenovirus is of human origin.
9. The method according to claim 1, wherein the first adenovirus is of simian origin.
10. The method according to claim 1, further comprising the step of isolating the chimeric adenovirus.
11. A method for generating a chimeric adenovirus for growth in a selected host cell, said chimeric adenovirus being derived from a parental adenovirus strain incapable of efficient growth in said host cell, said method comprising the step of generating a chimeric adenovirus comprising:
5' and 3' terminal regions of a first adenovirus which grows in a selected host cell type, said 5' terminal regions comprising the 5' inverted terminal repeat (ITRs) and necessary E1 gene functions, and said 3' terminal regions comprising inverted terminal repeat (ITRs) and necessary E4 gene functions; and internal regions from a parental adenovirus which lacks its native 5' and 3' terminal regions, said internal regions comprising the hexon, penton base and fiber;

wherein the resulting chimeric adenovirus comprises, from 5' to 3', the 5' terminal region of the first adenovirus, the internal region of the parental adenovirus, and the 3' terminal regions of the first adenovirus.

12. A chimeric adenovirus produced according to the method of claim 1.

13. A chimeric adenovirus comprising a hexon protein of a selected adenovirus serotype which is incapable of efficient growth in a selected host cell, said modified adenovirus comprising:

(a) adenovirus sequences of the left terminal end of a first adenovirus which grows in a selected host cell type, said left end region comprising the E1a, E1b and 5' inverted terminal repeat (ITRs);

(b) adenovirus sequences of the internal region of the selected adenovirus serotype which is incapable of efficient growth in the selected host cell, said internal region comprising the genes encoding the penton, hexon and fiber of the selected adenovirus;

(c) adenovirus sequences of the right terminal end of the first adenovirus, said right end region comprising the necessary E4 gene functions and the 3' inverted terminal repeat (ITRs),

wherein the resulting chimeric adenovirus comprises adenoviral structural and regulatory proteins necessary for infection and replication.

14. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus further comprises the IIIa, 52/55kDa and terminal protein (pTP) of the selected adenovirus serotype.

15. The chimeric adenovirus according to claim 13, wherein chimeric adenovirus comprises the polymerase of the first adenovirus.

16. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus expresses a functional chimeric protein formed from the first adenovirus and the selected adenovirus, said chimeric protein is selected from the group consisting of polymerase, terminal protein, 52/55 kDa protein, and IIIa.

17. The chimeric adenovirus according to claim 13, wherein the chimeric adenovirus comprises the terminal protein, 52/55 kDa, and/or IIIa of the selected adenovirus.

18. A host cell comprising a chimeric adenovirus according to claim 12.

19. The host cell according to claim 18, wherein said host cell is a human cell.

20. An isolated simian adenovirus nucleic acid sequence selected from the group consisting of:

(a) SA18 having the sequence of nucleic acids 1 to 31967 of SEQ ID NO:12 and

(b) a nucleic acid sequence complementary to the sequence of any of (a) to (f).

21. An isolated simian adenovirus serotype nucleic acid sequence selected from one or more of the group consisting of:

(a) 5' inverted terminal repeat (ITR) sequences;

(b) the adenovirus E1a region, or a fragment thereof selected from among the 13S, 12S and 9S regions;

(c) the adenovirus E1b region, or a fragment thereof selected from among the group consisting of the small T, large T, IX, and IVa2 regions;

(d) the E2b region;

(e) the L1 region, or a fragment thereof selected from among the group consisting of the 28.1 kD protein, polymerase, agnoprotein, 52/55 kD protein, and IIIa protein;

(f) the L2 region, or a fragment thereof selected from the group consisting of the penton, VII, VI, and Mu proteins;

(g) the L3 region, or a fragment thereof selected from the group consisting of the VI, hexon, or endoprotease;

(h) the 2a protein;

- (i) the L4 region, or a fragment thereof selected from the group consisting of the 100 kD protein, the 33 kD homolog, and VIII;
- (j) the E3 region, or a fragment thereof selected from the group consisting of E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, and E3 ORF9;
- (k) the L5 region, or a fragment thereof selected from a fiber protein;
- (l) the E4 region, or a fragment thereof selected from the group consisting of E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2, and E4 ORF1; and
- (m) the 3' ITR, of any of SA18 SEQ ID NO:12, or a sequence complementary to any of (a) to (m).

22. A simian adenovirus protein encoded by the nucleic acid sequence according to claim 21.
23. A composition comprising a simian adenovirus capsid protein according to claim 22 linked to a heterologous molecule for delivery to a selected host cell.
24. A method for targeting a cell having an adenoviral receptor comprising delivering to a subject a composition according to claim 23.
25. A nucleic acid molecule comprising a heterologous simian adenoviral sequence according to claim 21.
26. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence encodes an adenoviral gene product and is operatively linked to regulatory control sequences which direct expression of the adenoviral gene product in a host cells.
27. The nucleic acid molecule according to claim 25, wherein said simian adenoviral sequence comprises the E1a region of SA18 SEQ ID NO:12.

28. A pharmaceutical composition comprising the nucleic acid molecule according to claim 27 and a physiologically compatible carrier.

29. A recombinant adenovirus having a capsid comprising a protein selected from the group consisting of:

- (a) a hexon protein of SA18, SEQ ID NO 13, or a unique fragment thereof;
- (b) a penton protein of SA18, SEQ ID NO: 14, or a unique fragment thereof;
- (c) a fiber protein of SA18, SEQ ID NO: 15, or a unique fragment thereof.

30. The recombinant adenovirus according to claim 29, wherein the capsid is of an artificial serotype.

31. The recombinant adenovirus according to claim 29, wherein said virus further comprises a heterologous gene operatively linked to sequences which direct expression of said gene in a host cell.

32. The recombinant adenovirus according to claim 29, further comprising 5' and 3' adenovirus cis-elements necessary for replication and encapsidation.

33. The recombinant adenovirus according to claim 29, wherein said vector lacks all or a part of the E1 gene.

34. A host cell comprising a heterologous nucleic acid molecule comprising the nucleic acid sequence according to claim 21.

35. The host cell according to claim 34, wherein said host cell is stably transformed with the nucleic acid molecule.

36. The host cell according to claim 34, wherein said host cell expresses one or more adenoviral gene products from said nucleic acid molecule, said adenoviral gene products selected from the group consisting of E1a, E1b, E2a, and E4 ORF6.

37. The host cell according to claim 34, wherein said host cell is stably transformed with a nucleic acid molecule comprising the simian adenovirus inverted terminal repeats.

38. A composition comprising a recombinant virus according to claim 29 in a pharmaceutically acceptable carrier.

39. A method for delivering a heterologous gene to a mammalian cell comprising introducing into said cell an effective amount of the recombinant virus according to claim 29.

40. A method for repeat administration of a heterologous gene to a mammal comprising the steps of:

- (a) introducing into said mammal a first vector which comprises the heterologous gene and
- (b) introducing into said mammal a second vector which comprises the heterologous gene;

wherein at least the first virus or the second vector is a virus according to claim 29 and wherein the first and second recombinant vector are different.

41. A method for producing a selected gene product comprising infecting a mammalian cell with the recombinant virus according to claim 29, culturing said cell under suitable conditions and recovering from said cell culture the expressed gene product.

42. A method for eliciting an immune response in a mammalian host against an infective agent comprising administering to said host an effective amount of the

recombinant adenovirus of claim 29, wherein said heterologous gene encodes an antigen of the infective agent.

43. The method according to claim 42, comprising the step of priming the host with a DNA vaccine comprising the heterologous gene prior to administering the recombinant adenovirus.

ABSTRACT OF THE DISCLOSURE

A method for providing an adenovirus from a serotype which does not grow
5 efficiently in a desired cell line with the ability to grow in that cell line is described. The method involves replacing the left and right termini of the adenovirus with the corresponding termini from an adenovirus which grow efficiently in the desired cell line. At a minimum, the left terminus spans the 5' inverted terminal repeat, the left terminus spans the E4 region and the 3' inverted terminal repeat. The resulting
10 chimeric adenovirus contains the internal regions spanning the genes encoding the penton, hexon and fiber from the serotype which does not grow efficiently in the desired cell. Also provided are vectors constructed from novel simian adenovirus sequences and proteins, host cells containing same, and uses thereof.

Fig. 1

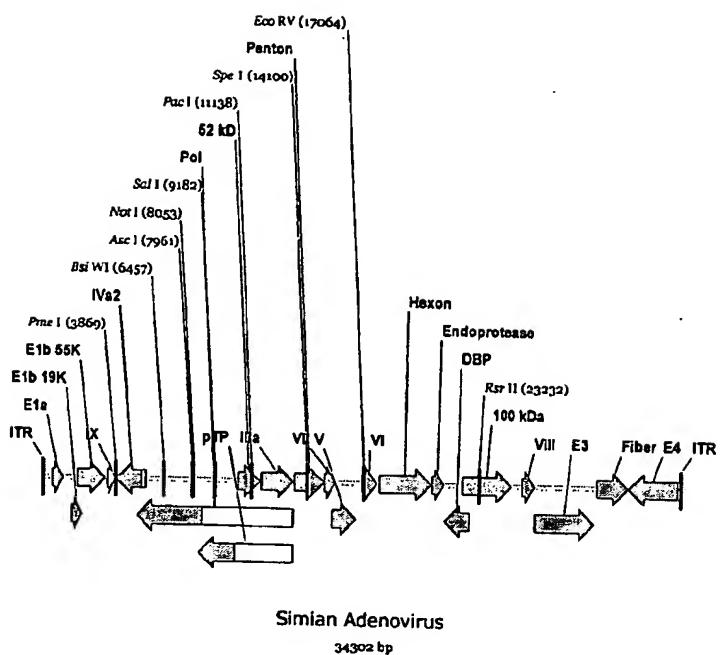
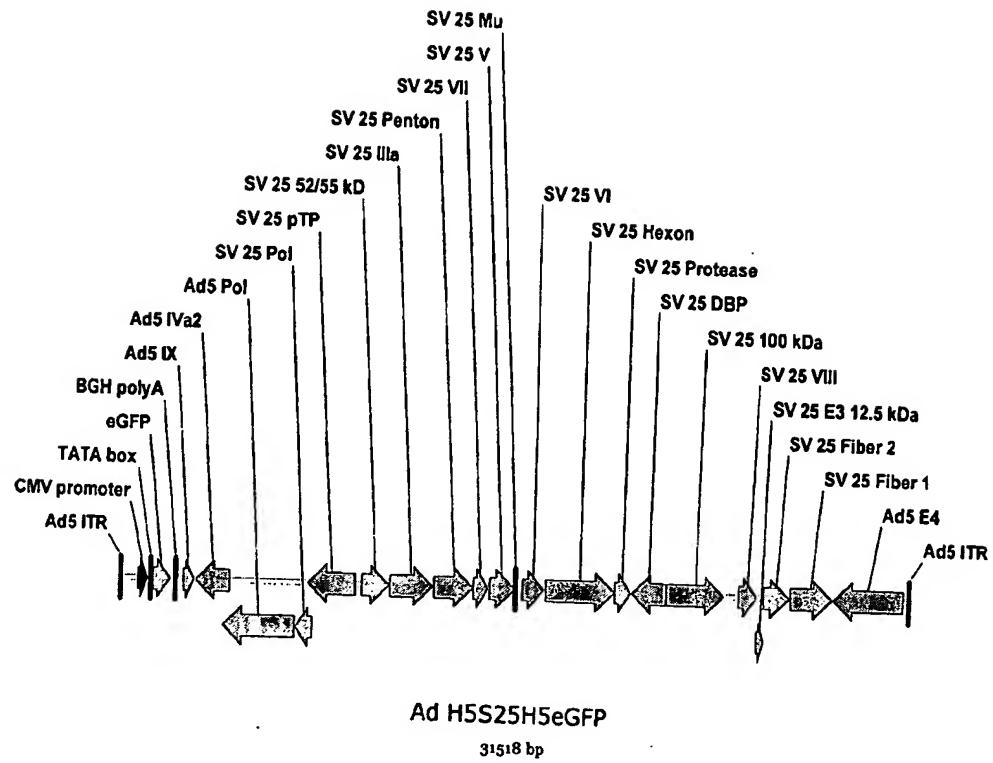


Fig. 2



APPLICATION DATA SHEET

Application Information	
Application Number::	
Filing Date::	May 28, 2004
Application Type::	Provisional
Subject Matter::	Utility
Suggested Classification::	
Suggested Group Art Unit::	
CD-ROM or CD-R::	None
Number of CD disks::	
Number of Copies of CDs::	
Sequence Submission?::	Yes
Computer Readable Form (CRF)?::	PAPER
Number of Copies of CRF::	
Title::	Methods of Generating Chimeric Adenoviruses and Uses for Such Chimeric Adenoviruses
Attorney Docket Number::	UPN-P3067-P3
Request for Early Publication?	No
Request for Non-Publication?	No
Suggested Drawing Figure::	
Total Drawing Sheets::	2
Small Entity::	No
Latin name::	
Variety denomination name	
Petition Included::	No
Petition Type	
Licensed US Govt. Agency::	
Contract or Grant Number::	
Secrecy Order in Parent Application::	

Applicant Information	
Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	Soumitra
Middle Name::	
Family Name::	Roy
Name Suffix::	
City of Residence::	Wayne
State or Province of Residence::	PA
Country of Residence::	US
Street of Mailing Address::	240 Pugh Road
City of Mailing Address::	Wayne
State or Province of Mailing Address::	PA
Country of Mailing Address::	US
Postal or Zip Code of Mailing Address::	19087

Applicant Information	
Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	James
Middle Name::	M.
Family Name::	Wilson
Name Suffix::	
City of Residence::	Gladwyne
State or Province of Residence::	PA
Country of Residence::	US
Street of Mailing Address::	1350 N. Avignon Drive
City of Mailing Address::	Gladwyne
State or Province of Mailing Address::	PA
Country of Mailing Address::	US
Postal or Zip Code of Mailing Address::	19035

Correspondence Information	
Correspondence Customer Number::	00270
Name::	Howson and Howson
Street of Mailing Address	Spring House Corporate Center, Box 457
City of Mailing Address	Spring House
State or Province of Mailing Address	Pennsylvania
Country of Mailing Address	US
Postal or Zip Code of Mailing Address::	19477
Phone Number::	215-540-9200
Fax Number::	215-540-5818
E-Mail Address::	ckodroff@howsonandhowson.com

Representative Information

Representative Customer No. 00270	Registration Number	Name
--------------------------------------	---------------------	------

Domestic Priority Information

Application	Continuity Type	Parent Application	Parent Filing Date

Foreign Priority Information

Country	Application Number	Filing Date	Priority Claimed

SEQUENCE LISTING

<110> Wilson, James M.
Roy, Soumitra
<120> Methods of Generating Chimeric Adenoviruses and Uses For Such
Chimeric Adenoviruses
<130> UPN-P3067
<160> 15
<170> PatentIn version 3.2
<210> 1
<211> 36462
<212> DNA
<213> chimpanzee adenovirus serotype Pan5
<400> 1
catcatcaat aatacacctc aaacttttgg tgcgcgtaa tatgcaaatg aggtatttga 60
atttggggat gcggggcggt gattggctgc gggagcgcc accgttaggg gcggggcggt
tgacgttttgc atgacgtggc cgtgaggcg agccggtttgc caagttctcg tggaaaagt 120
gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa tttcccgcc ctctctgaca 180
ggaaatgagg tggatcgatcag tgaaaacggg ccattttcgc gcgaaaactg 240
aatgaggaag tgaaaatctg agtaattccg cgtttatggc agggaggagt atttgcgag 300
ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatccc tcacctaatt 360
ttccgcgtac ggtgtcaaag tccgggtttt ttacgttagt gtcagctgat cgccagggt 420
tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagtttct 480
ccctccgcgc gcgagtcaaga tctacacttt gaaagatgag gcacctgaga gacctgcccg 540
gtaatgtttt cctggctact gggAACGAGA ttctggaaact ggtgggtggac gccatgatgg 600
gtgacgaccc tccggagccc cctaccccat ttgaagcgcc ttcgctgtac gatttgtatg 660
atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcgggtgaat gatttgttta 720
gcgtgcgcgc gctgtggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 780
ccctctctcca taccggaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 840
aagagctcga cctgcgtgc tatgaggaat gcttgcctcc gagcgatgat gaggaggacg 900
ggaggcgat tcgagctgca gcaaccagg gatgtgaaaaac agcgagcgag ggctttagcc 960
ggactgtcc tactctgccc ggacacggct gtaagtcttgc tgaatttcat cgcatgaata 1020
tggagataa gaatgtgatg tggccctgt gctatatgag agcttacaac catttgtttt 1080
1140

acagtaagt	tgattaactt	tagctgggaa	ggcagagggt	gactgggtgc	tgactggttt	1200
atttatgtat	atgttttta	tgttaggtc	ccgtctctga	cgtagatgag	acccccacta	1260
cagagtgcac	ttcatcaccc	ccagaaattt	gcgaggaacc	gcccgaagat	attattcata	1320
gaccagttgc	agtgagagtc	accgggcgt	a g c a g t g t g	ggagagttt	gatgacttgc	1380
tacagggtgg	ggatgaacct	ttggacttgt	gtacccggaa	acgccccagg	cactaagtgc	1440
cacacatgtg	tgtttactta	aggtgatgtc	agatattata	gggtgtggag	tgcaataaaa	1500
tccgtgttga	ctttaagtgc	gtggttatg	atccagggtt	ggggactgtg	ggtatataag	1560
caggtgcaga	cctgtgtgg	cagttcagag	caggactcat	ggagatctgg	acagtcttgg	1620
aagactttca	ccagactaga	cagctgttag	agaactcatc	ggagggagtc	tcttacctgt	1680
ggagattctg	tttcgggtgg	cctctagcta	agctagtcta	tagggccaag	caggattata	1740
aggatcaatt	tgaggatatt	ttgagagagt	gtcctggat	ttttgactct	ctcaacttgg	1800
gccatcagtc	tcactttaac	cagagtattc	tgagagccct	tgactttct	actcctggca	1860
gaactaccgc	cgcggtagcc	tttttgcct	ttatccttga	caa atggagt	caagaaaccc	1920
atttcagcag	ggattaccgt	ctggactgct	tagcagtagc	tttgtggaga	acatggaggt	1980
gccagcgct	gaatgcaatc	tccggctact	tgccagtaca	gccggtagac	acgctgagga	2040
tcctgagtct	ccagtcaccc	caggaacacc	aacgcgcaca	gcagccgcag	caggagcagc	2100
agcaagagga	ggaccgagaa	gagaacctga	gagccggct	ggaccctccg	gtggcggagg	2160
aggaggagta	gctgacttgt	ttccc gagct	gcgcgggtg	ctgacttagt	cttccagtg	2220
acggggagagg	gggattaagc	gggagaggca	tgaggagact	agccacagaa	ctgaactgac	2280
tgtcagtctg	atgagtgcac	ggcgccaga	atcggtgtgg	tggcatgagg	tgca	2340
ggggatagat	gaggctcag	tgtatgcata	gaaatattcc	ctagaacaag	tcaagacttg	2400
ttgggtggag	cccgaggatg	attgggaggt	agccatcagg	aattatgcac	agctggctct	2460
gaggccagac	aagaagtaca	agattacca	actgattaat	atcagaattt	cctgctacat	2520
ttcaggaaat	ggggccgagg	tggagatcag	tacccaggag	agggtggcct	tcagatgt	2580
catgatgaat	atgtacccgg	gggtgggtgg	catggaggga	gtcaccttta	tgaacgcgag	2640
gttcagggtt	gatgggtata	atgggggtgg	ctttatggcc	aacaccaagc	tgacagtgc	2700
cgatgctcc	ttctttggct	tcaataacat	gtgcattgag	gcctggggca	gtgttcagt	2760
gaggggatgc	agtttttcag	ccaaactggat	gggggtcg	ggcagaacca	agagcatggt	2820

tgtggtcctg	ggtgagctcg	tcataggcca	ttttaatgaa	tttggggcgg	agggtgcccg	4620
actgggggac	gaaggtgccc	tcgatcccgg	gggcgttagtt	gccctcgca	atctgcac	4680
cccaggcctt	gagctcgag	ggggggatca	tgtccacctg	cggggcgatg	aaaaaaacgg	4740
tttccggggc	ggggggatg	agctggccg	aaagcaggtt	ccggagcagc	tggacttgc	4800
cgcagccgt	ggggccgtag	atgaccggca	tgaccggctg	caggtggtag	ttgagggaga	4860
gacagctgcc	gtcctcgccg	aggaggggg	ccacctcg	catcatctcg	cgcacatgca	4920
tgttctcgcg	cacgagtcc	gccaggaggc	gctcgcccc	aagcgagagg	agctcttgca	4980
gcgaggcga	gttttcagc	ggcttgagcc	cgtcgccat	ggcattttg	gagagggtct	5040
gttgcagag	ttccagacgg	tcccagagct	cggtgatgtg	ctctaggca	tctcgatcca	5100
gcagacctcc	tcgttgcg	ggttggggcg	actgcgggag	tagggcacca	ggcgatggc	5160
gtccagcga	gccagggtcc	ggtccttcca	ggggcgcagg	gtccgcgtca	gcgtggctc	5220
cgtcacggtg	aagggggtcg	cgccgggctg	ggcgcttgcg	agggtgcgct	tcaaggctcat	5280
ccggctggtc	gagaaccgct	cccggtcg	gcctgcgcg	tccggccagg	agcaattgag	5340
catgagttcg	tagttgagcg	cctcgccgc	gtggcccttg	gcgcggagct	tacctttgga	5400
agtgtgtccg	cagacgggac	agaggaggga	cttggggcg	tagagcttgg	ggcgagggaa	5460
gacggactcg	ggggcgtagg	cgtccgcgc	gcagctggcg	cagacggct	cgcactccac	5520
gagccaggtg	aggctggcc	ggtcggggtc	aaaaacgagg	tttccctccgt	gtttttgtat	5580
gcgtttcta	cctctggct	ccatgagctc	gtgtccccgc	tgggtgacaa	agaggctgtc	5640
cgtgtccccg	tagaccgact	ttatggccg	gtcctcgagc	gggggtgcgc	ggtcctcg	5700
gtagaggaac	cccgcact	ccgagacgaa	ggccgggtc	caggccagca	cgaaggaggc	5760
cacgtggag	gggttagcggt	cgttgtccac	cagcgggtcc	accttctcca	gggtatgcaa	5820
gcacatgtcc	ccctcg	catccaggaa	ggtgattggc	ttgttaagtgt	aggccacgtg	5880
accgggggtc	ccggccgggg	gggtataaaa	ggggggggc	ccctgctcg	cctca	5940
ttccggatcg	ctgtccagga	gcgcagctg	ttggggtagg	tattccctct	cgaaggcggg	6000
catgacctcg	gcactcagg	tgtcagttc	tagaaacgag	gaggattga	tattgacgg	6060
gccgtggag	acgccttca	tgagccctc	gtccatctgg	tcaaaaaaga	cgatctttt	6120
gttgcagc	ttgggtggcga	aggagccgt	gagggcg	gagagcagct	tggcgatgg	6180
gcgcattggc	tggttttt	ccttgcggc	gcgccttgc	gcggcgatgt	tgagctgcac	6240

gtactcgccgc	gccacgcact	tccattcggg	gaagacggtg	tgagcttgc	cgccacat	6300
tctgacccgc	cagccgcgt	tgtgcagggt	gatgagggtcc	acgctggtgg	ccacctcgcc	6360
gcgcaggggc	tcgttggtcc	agcagaggcg	cccgccttg	cgcgacaga	agggggggcag	6420
cgggtccagc	atgagctcg	cgggggggtc	ggcgccacg	gtgaagatgc	cgggcaggag	6480
ctcggttgtcg	aagttagctga	tgcaggtgcc	cagatcgcc	agcgcgcctt	gccagtcgag	6540
cacggccagc	gcgcgcctcg	aggggctgag	gggcgtgcc	cagggcatgg	ggtgcgtgag	6600
cgcggaggcg	tacatgccgc	agatgtcgta	gacgttagagg	ggctcctcga	ggacgcccgt	6660
gtaggtgggg	tagcagegcc	ccccgcggat	gctggcgcc	acgtagtcgt	acagtcgtg	6720
c gagggcgcg	aggagcccg	tgcgcgggtt	ggagcgctgc	ggctttcgg	cgcggtagac	6780
gatctggcgg	aagatggcgt	gggagttgga	ggagatggtg	ggcctctgga	agatgttcaa	6840
gtgggcgtgg	ggcagtccga	ccgagtcct	gatgaagtgg	gcgttaggag	cctgcagctt	6900
ggcgaacgac	tcggcggtga	cgaggacgtc	cagggcgcag	tagtcgaggg	tctttggat	6960
gatgtcgta	ttgagctggc	ccttcgtt	ccacagctcg	cgggtgagaa	ggaactcttc	7020
gcggccttc	cagtacttt	cgagggggaa	ccgtcctga	tcggcacgg	aagagccac	7080
catgtagaac	tggttgacgg	cctttaggc	gcagcagccc	ttctccacgg	ggagggcgta	7140
agcttgcgcg	gccttgcgc	gggaggtgt	ggtgaggcgc	aaggtgtcg	gcaccatgac	7200
ctttaggaac	tggtgcttga	agtcgaggc	gtcgacgc	ccctgtccc	agagctggaa	7260
gtccgtgcgc	ttttttagg	cggggttggg	caaagcgaaa	gtaacatgt	tgaagaggat	7320
cttgcggcg	cggggcatga	agttgcgagt	gatgcggaaa	ggctggggca	cctcgccccg	7380
gttggatg	acctgggcgg	cgaggacgt	ctcgctgaag	ccgttgcgt	tgtccccgac	7440
gatgtagat	tccacgaatc	gcggcgcc	ttgacgtgg	ggcagcttct	tgagctcg	7500
gtaggtgagc	tcggcggtgt	cgctgaggcc	gtgctgcgt	agggcccagt	cggcgagggt	7560
gggggtggcg	ccgaggaagg	aagtccagag	atccacggcc	agggcggtct	gcaagcggtc	7620
ccggtaactga	cggaactgct	ggcccacggc	catttttgc	gggggtgacgc	agtagaaggt	7680
gcgggggtcg	ccgtgccagc	ggtcccactt	gagctggagg	gcgaggctgt	ggcgagctc	7740
gacgagcggc	gggtccccgg	agatttcat	gaccagcatg	aaggggacga	gctgcttgc	7800
gaaggacccc	atccaggtgt	aggtttcac	gtcgtaggtg	aggaagagcc	tttcgggtcg	7860
aggatgcgag	ccgatgggaa	agaactggat	ctcctgcac	cagttggagg	aatggctgtt	7920
gatgtatgg	aagttagaaat	gcccacggcg	cgccgagcac	tcgtgtttgt	gtttatacaa	7980

gcgtccgcag	tgctcgcaac	gctgcacggg	atgcacgtgc	tgcacgagct	gtacacctgggt	8040
tcctttgacg	aggaatttca	gtgggcagtg	gagcgctggc	ggctgcatct	ggtgctgtac	8100
tacgtccctgg	ccatcggtgt	ggccatcgtc	tgcctcgatg	gtggtcatgc	tgacgaggc	8160
gcccgggagg	caggccaga	cctcggtcg	gacgggtcg	agagcgagga	cgagggcg	8220
caggccggag	ctgtccaggg	tcctgagacg	ctgcggagtc	aggtcagtgg	gcagcggc	8280
cgccgggttg	acttgcagga	gctttccag	ggcgccggg	aggtccagat	ggtacttgat	8340
ctccacggcg	ccgttggtgg	cgacgtccac	ggcttgcagg	gtccctgtcc	cctggggcgc	8400
caccacccgtg	ccccgttct	tcttgggtgc	tggcgccggc	ggctccatgc	ttagaagcgg	8460
cggcgaggac	gcgcgcgggg	cggcaggggc	ggctcggggc	ccggaggcag	ggcgccagg	8520
ggcacgtcg	cgccgcgcgc	gggcagggttc	tggtaactgcg	cccgagaaag	actggcgtga	8580
gcgacgacgc	gacgggttgc	gtccctggatc	tgacgcctct	gggtgaaggc	cacgggaccc	8640
gtgagtttga	acctgaaaga	gagttcgaca	aatcaatct	cggtatcg	gacggcggcc	8700
tgccgcagga	tctcttgac	gtcgcccgag	ttgtccctgg	aggcgatctc	ggtcatgaac	8760
tgctcgatct	cctcttcctg	aaggctccg	cgaccggcgc	gctcgacggt	ggccgcgagg	8820
tctttggaga	tgcggccat	gagctgcag	aaggcg	tgcggccctc	gttccagacg	8880
cggctgtaga	ccacggctcc	gtcggttcg	cgcgccgc	tgaccacctg	ggcgagggtt	8940
agctcgacgt	ggcgctgaa	gaccgcgtag	ttgcagaggc	gctggtagag	gtagttgagc	9000
gtggtggcga	tgtgctcggt	gacgaagaag	tacatgatcc	agcggcggag	cgccatctcg	9060
ctgacgtcgc	ccagggcttc	caagcgctcc	atggcctcg	agaagtccac	ggcgaagtt	9120
aaaaactggg	agttgcgcgc	cgagacggc	aactcttc	ccagaagacg	gatgagctcg	9180
gcatgggtgg	cgcgcaccc	gcatcgaa	gccccgggg	gtcccttcc	ttccatctcc	9240
tcctcctctt	ccatcttc	caactac	tcttctactt	ccctctcagg	aggcgccggc	9300
gggggagggg	ccctgcgtcg	ccggcgccgc	acgggcagac	ggtcgtgaa	gcatcgatg	9360
gtctccccgc	gcccggcgc	catgtctcg	gtgacggcgc	gccccgtctc	gccccggccgc	9420
agctgtaa	cgccgcgc	catctccagg	tggccgcgg	gggggtctcc	gttggcagg	9480
gagagggcgc	tgacgtatgca	tcttatcaat	tggcccgtag	ggactccgc	caaggacctg	9540
agctctcg	gatccacggg	atccaaaac	cgctgaacga	aggcttcgag	ccagtcgcag	9600
tcgcaaggta	ggctgagccc	ggtttcttgc	tcttcggta	tttggcggg	aggcgccgg	9660

gcgatgtgc	tggatgaa	gttgaagtag	gcggctctga	gacggcgat	ggtggcgagg	9720
agcaccagg	ccttggccc	ggcttgcgg	atgcgcagac	ggtcggccat	gccccaggcg	9780
tggctctgac	acctggcgag	gtccttgcgg	tagtctgcg	tgagccgcgc	cacggcacc	9840
tcctccctgc	ccgcgcggcc	gtgcgcgcgc	gtgagcccg	acccgcgcgt	cggctggacg	9900
agcgcagg	cggcgacgac	gcgcgcggcg	aggatggcct	gctggatctg	ggtgagggtg	9960
gtctggaa	gtcgaagtc	gacgaagcgg	tggtaggctc	cggtggtat	ggtgtaggag	10020
cagttggcca	tgacggacca	gttgcggcgc	tggtgccgg	ggcgcacgag	ctcggtgtac	10080
ttgaggcg	agtaggcg	cgtgtcgaag	atgtagtcgt	tgcaggtgcg	cacgaggtac	10140
tggatccga	cgaggaagt	cggcgccggc	tggcggtaga	gccccatcg	ctcggtggcg	10200
ggggcgccgg	gcgcgagg	ctcgagcatg	aggcggtgg	agccgttagat	gtacctggac	10260
atccaggta	tgcggcg	ggtggtgag	gcgcgcggga	actcgccgac	gcgggtccag	10320
atgttgcgc	gcggcaggaa	gtagttcatg	gtggccgcgg	tctggccgt	gaggcg	10380
cagtcgtg	tgctctagac	atacggcaa	aaacgaaagc	ggtcagcggc	tgcactccgt	10440
ggcctggagg	ctaagcgaac	gggttggct	gcgcgtgtac	cccggttgc	gtccctgc	10500
gaatcagg	ggagccgcag	ctaacgtgg	actggcactc	ccgtctcgac	ccaagcctgc	10560
taacgaaacc	tccaggatac	ggaggcg	cgttttggcc	attttcgta	ggccggaaat	10620
gaaactagta	agcgcggaaa	gccccgtcc	gcgcgtgc	gctgccgt	tctggagaaa	10680
gaatcgccag	gttgcgttg	cggtgtgccc	cggttcgagc	ctcagcgc	ggcgccggcc	10740
ggattccgc	gtaacgtgg	gcgtggctc	cccgctgttt	ccaagacccc	ttagccagcc	10800
gacttctcca	gttacggagc	gagccctct	ttttctgtg	ttttgc	atgcatccc	10860
tactgcggca	gatgcgcccc	caccctccac	cacaaccgc	cctaccgc	cagcagcaac	10920
agccggcg	tctgcccc	ccccagc	agcagccagc	cactaccgc	gcggccggc	10980
tgagcggagc	cgccgttc	atgacctgg	ccttggaa	ggcgaggggg	ctggcgccgc	11040
tggggcg	gtcgccggag	cgccacccgc	gcgtgcagat	aaaaagg	gctcg	11100
cctacgtgc	caagcagaac	ctgttgcag	acaggagcgg	cgaggagcc	gaggagatgc	11160
gcgcctcccg	cttccacgc	gggcgggagc	tgccggcg	cctggaccga	aagcgggtgc	11220
tgagggacga	ggatttc	gagcggacg	tgacggggat	cagccccgc	cgcgcgcacg	11280
tggccgcggc	caacctggc	acggcgta	agcagaccgt	gaaggaggag	agcaacttcc	11340
aaaaatcctt	caacaaccac	gtgcgcacgc	tgcgcgc	cgaggaggtg	accctggcc	11400

tgatgcacct gtgggacctg ctggaggcca tcgtgcagaa ccccacgagc aagccgctga 11460
cggcgccagct gtttctggtg gtgcagcaca gtcgggacaa cgagacgttc agggaggcgc 11520
tgctgaatat caccgagccc gagggccgct ggctcctgga cctggtgaac attctgcaga 11580
gcatcgtggt gcaggagcgc gggctgccgc tgccgagaa gctggcggcc atcaacttct 11640
cggtgcgtgag cctgggcaag tactacgcta ggaagatcta caagaccccg tacgtgcccc 11700
tagacaagga ggtgaagatc gacgggtttt acatgcgcac gaccctgaaa gtgctgaccc 11760
tgagcgcacga tctgggggtg taccgcaacg acaggatgca ccgcgcggtg agcgcagcc 11820
gccggcgccga gctgagcgcac caggagctga tgcacagcct gcagcgggccc ctgaccgggg 11880
ccgggaccga gggggagagc tactttgaca tggcgccgga cctgcgtgg cagcttagcc 11940
gccgggcctt ggaagctgcc ggcgggttccc cctacgtgga ggaggtggac gatgaggagg 12000
aggagggcga gtacctggaa gactgatggc gcgaccgtat ttttgcata gtcagcaaca 12060
gccacccgccc cctcctgatc ccgcgcgtgc ggcggcgctg cagagccagc cgtccggcat 12120
taactcctcg gacgattgga cccaggccat gcaacgcacatc atggcgctga cgacccgcaa 12180
tcccgaagcc tttagacagc agcctcaggc caaccgactc tcggccatcc tggaggccgt 12240
ggtgtccctcg cgctcgaacc ccacgcacga gaaggtgctg gccatcgtga acgcgtgg 12300
ggagaacaag gccatccgcg ggcacgagc cggctggtg tacaacgcgc tgctggagcg 12360
cgtggcccgca tacaacagca ccaacgtgca gacgacccgt gaccgcatttg tgaccgacgt 12420
gcgcgaggcg gtgtcgccgc ggcgcgggtt ccaccgcgcag tcgaacctgg gctccatgg 12480
ggcgctgaac gccttcctga gcacgcagcc cgccaaacgtg ccccgccggc aggaggacta 12540
caccaacttc atcagcgcgc tgcggctgat ggtggccgag gtgccccaga gcgaggtgta 12600
ccagtccggg ccggactact tttccagac cagtcgcccag ggcttgcaga ccgtgaacct 12660
gagccaggct ttcaagaact tgcagggact gtggggcgtg caggccccgg tcggggaccg 12720
cgcgacggtg tcgagccctgc tgacgcccggaa ctcgcgcctg ctgctgctgc tggggccgc 12780
cttcacggac agcggcagcg tgagccgcga ctcgtacctg ggctacctgc ttaacctgta 12840
ccgcgaggcc atcgggcagg cgacgtgga cgagcagacc taccaggaga tcacccacgt 12900
gagccgcgcg ctggggcagg aggacccggg caacctggag gccaccctga acttcctgct 12960
gaccaaccgg tcgcagaaga tccccccca gtacgcgcgtg agcaccggagg aggagcgcac 13020
cctgcgtacat gtgcagcaga ggcgtggggct gttcctgatg caggaggggg ccacgcccag 13080

cgccgcgctc gacatgaccg cgcgcaacat ggagcccagc atgtacgccc gcaaccgccc 13140
gttcatcaat aagctgatgg actacttgca tcggggggcc gccatgaact cggactactt 13200
taccaacgcc atcttgaacc cgcactggct cccgccccccc gggttctaca cggggcgagta 13260
cgacatgccc gaccccaacg acgggttcct gtgggacgac gtggacagca gcgtgttctc 13320
gccgcgcccc accaccacca ccgtgtggaa gaaagagggc ggggaccggc ggccgtcctc 13380
ggcgctgtcc ggtcgcgccg gtgtcgccgc ggcggtgcggc gaggccgcca gccccttccc 13440
gagcctgccc ttttcgctga acagcgtgcg cagcagcggag ctgggtcgcc tgacgcggcc 13500
gcgcctgctg ggcgaggagg agtacctgaa cgactccttgc ttccggcccg agcgcgagaa 13560
gaacttcccc aataacggga tagagacgcgt ggtggacaag atgagccgct ggaagacgta 13620
cgcgcacgag cacagggacg agccccgagc tagcagcggc accggcgccca cccgttagacg 13680
ccagcggcac gacaggcagc ggggtctgggt gtgggacgat gaggattccg ccgacgcacag 13740
cagcgtgttgc gacttgggtg ggagtgggtgg tggtaaccccg ttcgctcacc tgccggcccg 13800
tatcgccgcgctc ctgtatgtaaatctgaaaaa aataaaaagac ggtactcacc aaggccatgg 13860
cgaccagcgt gcgttcttct ctgttgttttgc tagtagtatg atgaggcgcg tgtacccgg 13920
gggtccctcttccctcgtacg agagcgtgat gcagcaggcg gtggcggccgg cgatgcagcc 13980
cccgctggag ggccttacg tgcccccccg gtacctggcg cctacggagg ggccggaaacag 14040
cattcgttac tcggagctgg cacccttgcgatcgataccacc cgggttgtacc tgggtggacaa 14100
caagtcggcg gacatcgccct cgctgaacta ccagaacgac cacagcaact tcctgaccac 14160
cgtgggtgcag aacaacgatt tcaccccccac ggaggccage acccagacca tcaactttga 14220
cgagcgcctcg cgggtggccg gccagctgaa aaccatcatg cacaccaaca tgcccaacgt 14280
gaacgagttc atgtacagca acaagttcaa ggccgggtg atggtctcgcc gcaagacccc 14340
caacggggtc acagtaacag atggtagtca ggacgagctg acctacgagt ggggtggagtt 14400
tgagctgccc gagggcaact tctcggtgac catgaccatc gatctgatga acaacgccc 14460
catcgacaac tacttggcggtggggccgaaacgggtg ctggagagcg acatcgccgt 14520
gaagttcgac acgcgcact tccggctggg ctgggacccc gtgaccgagc tgggtatgcc 14580
gggcgtgtac accaacgagg cttccaccc cgacatcgatcgatccctgcgatcc 14640
ggacttcacc gagagccgccc tcagcaacct gctggcattc cgcaagcggc agcccttcca 14700
ggagggcttc cagatcctgt acgaggaccc ggaggggggc aacatcccccg cgctgtgg 14760
cgtggacgccc tacgagaaaa gcaaggagga tagcggccgccc gccggcgaccg cagccgtggc 14820

caccgcctct accgaggtgc gggcgataa ttttgc tagc gccgcgacac tggcagcggc 14880
cgaggcggct gaaaccgaaa gtaagatagt gatccagccg gtggagaagg acagaagga 14940
gaggagctac aacgtgctcg cggacaagaa aaacaccgcc taccgcagct ggtacctggc 15000
ctacaactac ggcgaccccg agaaggcgt gcgcctcgg acgctgccta ccacctcgga 15060
cgtcacctgc ggcgtggagc aagtctactg gtcgcgtccc gacatgtgc aagacccgg 15120
caccttccgc tccacgcgtc aagtttagcaa ctacccggtg gtggcgcggc agctcctgcc 15180
cgtctactcc aagagcttct tcaacgagca ggccgtctac tcgcagcagc tgcgccctt 15240
cacctcgctc acgcacgtct tcaaccgtt ccccgagaac cagatcctcg ttgcggccgc 15300
cgcgccccacc attaccaccc tcagtaaaa cgttcctgtct ctcacagatc acgggaccct 15360
gccgctgcgc agcagtatcc ggggagtcca gcgcgtgacc gtcactgacg ccagacgcgg 15420
cacctgcccc tacgtctaca aggcctggg cgtagtcgcg ccgcgcgtcc tctegagccg 15480
caccttctaa aaaatgtcca ttctcatctc gcccagtaat aacaccggtt gggcctgcg 15540
cgcgccccacc aagatgtacg gaggcgtcg ccaacgctcc acgcaacacc ccgtgcgcgt 15600
gcgcgggcac ttccgcgtc cctggggcgc cctcaaggcc cgctgcgtc cgccgaccac 15660
cgtcgacgac gtgatcgacc aggtggtggc cgacgcgcgc aactacaacgc ccgcgcgcgc 15720
gccegtctcc accgtggacg ccgtcatcg cagcgtggtg gccgacgcgc gcccgtacgc 15780
ccgcgcacaag agccggcggc ggccatcgcc cccggggcac cggagcaccc ccgcattgcg 15840
cgccggcgca gccttgcgtc gcaggggccag ggcacggga cgcaggggca tgctcagggc 15900
ggccagacgc gcggcctccg gcagcagcag cggccggcagg accccgcagac ggcggccac 15960
ggcggcggcg gcggccatcg ccagcatgtc cccggggcgcc cgccggcaacg tgtactgggt 16020
gcgcgacgcc gccaccggtg tgcgctgccc cgtgcgcacc cgcggccctc gcacttgaag 16080
atgctgactt cgcgatgtt atgtgtccca gcggcgagga ggatgtccaa gcgcatttc 16140
aaggaaagaga tgctccaggt catcgccct gagatctacg gcccggcggc ggtgaaggag 16200
gaaagaaagc cccgcaaaact gaagcgggtc aaaaaggaca aaaaggagga ggaagatgt 16260
gacggactgg tggagttgt gcgcgatgtc gccccccggc ggccgcgtc gtggcgcggg 16320
cgaaaagtga aaccgggtct gcgcacccggc accacggtg tcttcacgc cggcgagcgt 16380
tccggctccg cctccaagcg ctccatcgac gaggtgtacg gggacgagga catcctcgag 16440
caggcggccg aacgtctggg cgagtttgct tacggcaagc gcagccgccc cgccgccttg 16500

aaagaggagg cgggtgtccat cccgctggac cacggcaacc ccacgcccag cctgaagccg 16560
gtgaccctgc agcaggtgct gcctggtgcg gcgcgcgc ggggcttcaa gcgcgagggc 16620
ggcgaggatc tgtacccgac catgcagctg atggtgccca agcgccagaa gctggaggac 16680
gtgctggagc acatgaaggt ggaccccgag gtgcagcccg aggtcaaggt gcggcccatc 16740
aagcaggtgg cccccggccct gggcgtgcag accgtggaca tcaagatccc cacggagccc 16800
atggaaacgc agaccgagcc cgtgaagccc agcaccagca ccatggaggt gcagacggat 16860
ccctggatgc cggcacccggc ttccaccacc cgccgaagac gcaagtacgg cgccgcccagc 16920
ctgctgatgc ccaactacgc gtcgtatcct tccatcatcc ccacgcccgg ctaccgcggc 16980
acgcgcctct accgcggcta caccagcagc cgccgcccga agaccaccac ccgcgcgcgc 17040
cgtcgtcgca cccgcgcgag cagcaccgcg acttccgcgc ccgcgcgtt gcggagagtg 17100
taccgcagcg ggccgcgcgc tctgaccctg ccgcgcgcgc gctaccaccc gagcatcgcc 17160
attnaactac cgcctctac ttgcagatata ggcctcaca tgccgcctcc gcgtccccat 17220
tacgggctac cgaggaagaa agccgcgcgc tagaaggctg acggggaaacg ggctgcgtcg 17280
ccatcaccac cggccgcggc gcgcctacag caagcggtt gggggaggtt tcctgcccgc 17340
gctgatgcgc atcatcgccg cggcgatcgg ggcgatcccc ggcatacgctt ccgtggcggt 17400
gcaggcctct cagcgccact gagacacagc ttggaaaatt ttaataaaaa aatggactga 17460
cgctccttgtt cctgtgatgt gtgttttag atggaaagaca tcaattttc gtccctggca 17520
ccgcgcacacg gcacgcggcc gtttatggc acctggagcg acatcgccaa cagccaaactg 17580
aacggggcg cttcaattt gggcgatctc tggagcgccc ttaagaattt cgggtccacg 17640
ctcaaaacctt atggcaacaa ggcgttggaaac agcagcacag ggcaggcgct gaggaaaag 17700
ctgaaagagc agaacttcca gcagaaggtg gtcgtggcc tggcctcggtt catcaacggg 17760
gtggggacc tggccaaacca ggccgtgcag aaacagatca acagccgcct ggacgcggc 17820
ccgcgcgcgg ggtccgtgga gatggcccgatggaggagg agctgcctcc cctggacaag 17880
cgccgcgaca agcgaccgcg tcccgacgcg gaggagacgc tgctgacgca caccggacgag 17940
ccgcgcgcgtt acgaggaggc ggtggaaactg ggtctgccc ccacgcggcc cgtggcgctt 18000
ctggccacccg gggtgctgaa acccagcagc agcagcagcc agcccgccgac cctggacttg 18060
cctccgcctg ttccacagtg gctaagcccc tgccgcgggtt gggcgatcg 18120
tcgcgcgcggc cccgaggccg ccccgaggcg aactggcaga gcaactctgaa cagcatcg 18180
ggtctggag tgcagatgtt gaagcgccgc cgctgttattt aaaagacact gtgcgttta 18240

acttgcttgt ctgtgtgtat atgtatgtcc gccgaccaga aggaggagga agaggcgcgt 18300
cgccgagttt caagatggcc accccatcga tgctgccccca gtgggcgtac atgcacatcg 18360
ccggacagga cgcttcggag tacctgagtc cgggtctgggt gcagttcgcc cgccacag 18420
acacctactt cagtctgggg aacaagttt a ggaaccccac ggtggcgccc acgcacatcg 18480
tgaccaccga ccgcagccag cggctgacgc tgccgttcgt gcccgtggac cgccaggaca 18540
acacctactc gtacaaagtgc cgctacacgc tggccgtgggg cgacaaccgc gtgctggaca 18600
tggccagcac ctactttgac atccgcggcg tgctggatcg gggcccttagc ttcaaaccct 18660
actccggcac cgcttacaac agcctggctc ccaagggagc gccaacact tgccagtgg 18720
catataaaacgtc tgatgggtat actggtagac aaaaaaccta tacatatggaa aatgcgcctg 18780
tgcaaggcat tagtattaca aaagatggta ttcaacttgg aactgacact gatgatcagc 18840
ccatttatgc agataaaaact tatcaaccag agcctcaagt gggtgtatgc gaatggcatg 18900
acatcaactgg tactgtatgaa aaatatggag gcagagctct caaggcctgac accaaaatga 18960
agccctgcta tggttctttt gccaaggccta ccaataaaaga aggaggcag gcaaatgtga 19020
aaaccgaaac aggccgtacc aaagaatatg acattgacat ggcattcttc gataatcgaa 19080
gtgcagctgc ggctggcctg gcccagaaa ttgtttgtt tactgagaat gtggatctgg 19140
aaactccaga tactcatatt gtatacaagg cgggcacaga tgacagcagc tcttctatca 19200
atttgggtca gcagtcctatg cccaacacac ccaactacat tggctttaga gacaacttta 19260
tcggcctcat gtactacaac agcaactggca acatgggcgt gctggctggt caggeetccc 19320
agctgaatgc tgtggtggac ttgcaggaca gaaacactga actgtcctac cagctctgc 19380
ttgactctct gggcgacaga accaggattt tcagttatgtg gaatcaggcg gtggacagct 19440
atgaccccgaa tggcgccatt attgaaaatc acgggtgtggaa ggatgaactc cctaaactatt 19500
gcttcccccgtt ggatgctgtg ggtagaactg atacttacca gggaaattaag gccaatggtg 19560
ctgatcaaaccac cacctggacc aaagatgata ctgttaatga tgctaatgaa ttggcaagg 19620
gcaatccctt cgccatggag atcaacatcc aggccaacct gtggcggAAC ttcctctacg 19680
cgaacgtggc gctgtacctg cccgactcct acaagttacac gcccggcaac atcacgctgc 19740
cgaccaacac caacacacac gattacatga acggccgcgt ggtggcgccc tgcgtggatgg 19800
acgcctacat caacatcggtt ggcgcgtggcgt cgcgtggaccc catggacaac gtcaacccct 19860
tcaaccacca cgcacacgcg ggcctgcgt accgcctccat gtcctggc aacggggcgct 19920

acgtgccctt ccacatccag gtgccccaaa agttcttcgc catcaagagc ctcctgctcc 19980
 tgcccggtc ctacacctac gagtggaact tccgcaagga cgtcaacatg atcctgcaga 20040
 gctccctcg cgacgacctg cgacggacg gggctccat cgccctcacc agcatcaacc 20100
 tctacgccac cttcttcccc atggcgacaa acaccgcctc cacgctcgag gccatgctgc 20160
 gcaacgacac caacgaccag tccttcaacg actacctctc ggccggcaac atgccttacc 20220
 ccatcccgcc caacgcccacc aacgtgccc tctccatccc ctgcgcacac tggccggct 20280
 tcccgccatg gtccttcacg cgccctcaaga cccgcgagac gccctcgctc ggctccgggt 20340
 tcgaccccta cttcgctcac tcgggctcca tcccctaccc cgacggcacc ttctacctca 20400
 accacacctt caagaagggtc tccatcacct tcgactcctc cgtcagctgg cccggcaacg 20460
 accgcctctt gacgcccac gagttcgaaa tcaagcgcac cgtcgacgga gaggggtaca 20520
 acgtggccca gtgcaacatg accaaggact ggttcctggt ccagatgtg gcccactaca 20580
 acatcggcta ccagggcttc tacgtgccc agggctacaa ggaccgcatt tactccttct 20640
 tccgcaactt ccagcccatg agccgccagg tcgtggacga ggtcaactac aaggactacc 20700
 aggccgtcac cctggcttac cagcacaaca actcgggctt cgtcggctac ctgcgeccca 20760
 ccatgcgcac gggacagcccc tacccgcac actaccccta cccgctcatc ggcaagagcg 20820
 cctgcgcac cgtcaccatg aaaaagtcc tctgcgaccg ggtcatgtgg cgcatccct 20880
 tctccagcaa cttcatgtcc atgggcgcgc tcaccgaccc cggccagaac atgccttacg 20940
 ccaactccgc ccacgcgcta gacatgaatt tcgaagtcga ccccatggat gagtccaccc 21000
 ttctctatgt tgtcttcgaa gtcttcgacg tcgtccgagt gcaccagccc caccgcggcg 21060
 tcatcgaggc cgtctacctg cgacgcctt tctcgccgg caacgcccacc acctaagccc 21120
 cgctcttgc tcttgcaga tgacggctg tgccggctcc ggccgagcagg agctcagggc 21180
 catcctccgc gacctgggtc gcggggccctg cttectggc accttcgaca agcgttccc 21240
 gggattcatg gccccgcaca agctggctg cggccatcgta aacacggccg gccgcgagac 21300
 cgggggcgag cactggctgg cttcgctg gaaccgcgc tcccacaccc gtcacccctt 21360
 cgaccccttc gggttctcg acgagcgcct caagcagatc taccagttcg agtacgaggg 21420
 cctgctgcgc cgcagcgcacc tggccaccga ggaccgcgtgc gtcacccctgg aaaagtccac 21480
 ccagacccgtg cagggccgc gtcggccgc ctgcgggctc ttctgctgca tggccctgca 21540
 cgccttcgtg cactggcccg accgccccat ggacaagaac cccaccatga acttgcgtac 21600
 ggggggtgcac aacggcatgc tccagtcgc ccaggtggaa cccaccctgc gccgcacca 21660

ggaggcgctc taccgcttcc tcaacgccc ctccgcctac ttgcgtcccc accgcgcg 21720
catcgagaag gccaccgcct tcgaccgcat gaatcaagac atgtaaaccg tgtgttatg 21780
tgaatgcattt attcataata aacagcacat gttatgcc a cttttctga ggctctgact 21840
ttat tagaa atcgaagggg ttctgcccgc ttcggcgtg ccccgccggc agggatacgt 21900
tgccgaactg gtacttggc agccacttga actcgggat cagcagcttc ggcacggg 21960
ggtcggggaa cgagtcgctc cacagcttgc gcgtgagttg cagggcgccc agcaggtcg 22020
gcccggagat cttgaaatcg cagttggac ccgcgttctg cgccgggag ttgcgtaca 22080
cggggttgca gcacttggAAC accatcaggg ccgggtgctt cacgctcgcc agcaccgtcg 22140
cgtcggtgat gcctccacg tccagatct cggcgttggc catcccgaa ggggtcatct 22200
tgcaggctcg ccgccccatg ctggcacgc agccgggctt gtgggtgcaa tcgcagtgca 22260
gggggatcag catcatctgg gcctgctcgg agtcatgccc cgggtacatg gccttcatga 22320
aagcctccag ctggcggaag gcctgctcg cttgcgcgc ctcggtaag aagacccgc 22380
aggacttgc agagaactgg ttggtggcgc agccggcgctc gtgcacgcag cagcgcgcgt 22440
cgttgttggc cagctgcacc acgctgcgc cccagcggtt ctgggtgatc ttggcccggt 22500
cggggttctc cttcagcgcg cgctgcccgt ttcgcgtcgcc cacatccatc tcgatcggt 22560
gctccttctg gatcatcaccg gtcccgtgca ggcacgcag cttgcctcg gcctcggtgc 22620
acccgtgcag ccacagcgcg cagccggtgc actcccgat tttgtggcgc atctggag 22680
gcgagtgcac gaagccctgc aggaagcggc ccacatcgat ggtcagggtc ttgttgcgg 22740
tgaaggctcg cgggatgccc cggtgctctt cttcacata caggtggcag atgcggcggt 22800
acacctcgcc ctgctcgccg atcagctggg aggccggactt caggtgcctc tccacgcgtt 22860
accggccat cagcagcgctc atgacttcca tgccttctc ccaggccgag acgatcgca 22920
ggctcagggg gtttttacc gcccgttgc ttttagtgc cggcgctgag gtcagggggt 22980
cgtttcgtc cagggtctca aacactcgat tgccgtctt ctcgggtatc cgacgggggg 23040
gaaagctgaa gcccacggcc gcccgttgc cttggcctg ctttcgttcc tcgcgttctt 23100
ggctgatgtc ttgcaaaggc acatgcttgg ttttgcgggg ttttttttgg ggcggcagag 23160
gcccggccgg agacgtgctg ggcgagcgcg agtttcgtct caccacgact atttttttt 23220
cttggccgtc gtccgagacc acgcggcggtt aggcacgtcct cttctggggc agaggcggag 23280
gcgacgggtc ctcgcgggttc ggcggccggc tggcagagcc cttccggcgt tcgggggtgc 23340

gctcctggcg ggcgtgtct gactgacttc ctccgggcc ggccatttg 23400
 gagcaacaag catggagact cagccatgt cgccaacatc gccatctgc cccggccg 23460
 ccgacgagaa ccagcagcag aatgaaagct taaccggccc gccgcccage cccacctccg 23520
 acggccggcgc ggccccagac atgcaagaga tggaggaatc catcgagatt gacctggct 23580
 acgtgacgccc cgccggagcac gaggaggagc tggcagcgcg cttttcagcc cccggaaagaga 23640
 accaccaaga gcagccagag caggaagcag agagcgagca gcagcaggct gggctcgagc 23700
 atggcgacta cctgagcggg gcagaggacg tgctcatcaa gcatctggcc egccaatgca 23760
 tcatcgtaa ggacgcgctg ctgcaccccg ccgaggtgcc cctcagcgtg gggagctca 23820
 gcccgccta cgagcgcaac ctcttcgcg cgcgcgtgcc ccccaagcgc cagcccaacg 23880
 gcacctgcga gcccaccccg cgcctcaact tctacccggt cttegcggtg cccgaggccc 23940
 tggccaccta ccacctttt ttcaagaacc aaaggatccc cgtctctgc cgcgcctaacc 24000
 gcacccgcgc cgacgcctg ctcaacctgg gtccccggc cccgcctacct gatatgcct 24060
 ccttggaaaga gttcccaag atttcgagg gtctggcag cgacgagact cggccgcga 24120
 acgctctgca aggaagcggg gaggagcatg agcaccacag cgcctggtg gagttggaaag 24180
 ggcacaacgc gcgcctggcg gtgtcaagc gcacggcgtga gctgaccac ttcgcctacc 24240
 cggcgctcaa cctggccccc aaggcatga ggcgcgtcat ggaccaggtg ctcatcaagc 24300
 ggcctcgcc cctctcgat gaggacatgc aggaccccgaa gagctcggac gaggcaagc 24360
 ccgtggtcag cgacgagcag ctggcgcgt ggctgggagc gagtagcacc cccagagct 24420
 tggaaagagcg ggcacaagctc atgatggccg tggcctggt gaccgtggag ctggagtgtc 24480
 tgcgcgcgtt ctgcgcgac gcagagaccc tgcccaaggt cgaggagaac ctgcactacc 24540
 tcttcaggca cgggtttgtg cgcgcgcgt gcaagatctc caacgtggag ctgcaccaacc 24600
 tggcctccta catgggcata ctgcacgaga accgcctggg gcagaacgtg ctgcacaccca 24660
 ccctgcgcgg ggaggcccg cgcgcactaca tccgcgcactg cgtctacctg tacctctgcc 24720
 acacctggca gacgggcata ggcgtgtggc agcagtgcct ggaggagcag aacctgaaag 24780
 agctctgca gtcctgcag aagaacctga aggcctgtg gaccgggttc gacgagcgc 24840
 ccacccgcctc ggacctggcc gacctcatct tcccccgcgc cctgcggctg acgctgcgc 24900
 acggactgccc cgactttatg agtcaaagca tggcaaaaa ctttcgtct ttcatcctcg 24960
 aacgctccgg gatectgccc gccacctgtc cgcgcgtgcc ctggacttc gtggcgcgtga 25020
 cctccgcga gtggcccccgg cgcgcgtggc ggcactgcta cctgcgcgc ctggccaaact 25080

acctggccta ccactcgac gtgatcgagg acgtcagcgg cgagggtctg ctcgagtgcc 25140
 actgccgctg caacctctgc acgcccacc gtccttggc ctgcaacccc cagctgctga 25200
 gcgagaccca gatcatcgac accttcgagt tgcaaggccc cggcgagggc aaggggggtc 25260
 tgaaaactcac cccggggctg tggacctcgg cttacttgcg caagttcgtg cccgaggact 25320
 accatccctt cgagatcagg ttctacgagg accaatccca gccgccccaa gccgaactgt 25380
 cggectgcgt catcaccagg gggccatcc tggcccaatt gcaagccatc cagaaatccc 25440
 gccaagaatt tctgctgaaa aaggggccacg gggtctacct ggaccccccag accggagagg 25500
 agctcaaccc cagttcccc caggatcccc cgaggaagca gcaagaagct gaaagtggag 25560
 ctgccgcgc cggaggattt ggaggaagac tggagagca gtcagggcaga ggaggaggag 25620
 atggaagact gggacagcac tcaggcagag gaggacagcc tgcaagacag tctgaaagac 25680
 gaggtggagg aggaggcaga ggaagaagca gccgcccaca gaccgtcgct ctcggcggag 25740
 aaagcaagca gcacggatac catctccgt ccgggtcggg gtcgcggcga cccggcccac 25800
 agtaggtggg acgagacccgg ggcgttcccc aaccccacca cccagacccgg taagaaggag 25860
 cggcagggat acaagtcctg gcgccccac aaaaacgcca tcgttcctg cttgcaagcc 25920
 tgcggggca acatctccctt cacccggcgc tacctgctct tccaccgcgg ggtgaacttc 25980
 ccccgcaaca tcttgcatta ctaccgtcac ctccacagcc cttactactg tttccaagaa 26040
 gaggcagaaa cccagcagca gcagaaaacc agccggcagca gcagctagaa aatccacagc 26100
 ggcggcaggt ggactgagga tcgcagcgaa cgagccggcg cagacccggg agctgaggaa 26160
 ccggatcttt cccacccctct atgcctatctt ccagcagagt cggggccagg agcaggaact 26220
 gaaagtcaag aaccgttctc tgcgtcgct cacccgcagt tgtctgtatc acaagagcga 26280
 agaccaactt cagcgcactc tcgaggacgc cgaggtcttc ttcaacaagt actgcgcgt 26340
 cactttaaa gagtagcccc cgcccgccca cacacggaaa aaggcgggaa ttacgtcacc 26400
 acctgcgtcc ttgcggcgtc catcatcatg agcaaagaga ttcccacgac ttacatgtgg 26460
 agctaccagc cccagatggg cctggccgc ggcggccccc aggactactc cacccgcgt 26520
 aactggctca ggcggggcc cgcgatgtac tcacgggtga atgacatccg cggccgcga 26580
 aaccagatac tcctagaaca gtcagcgatc accgcccacgc cccgccatca ccttaatccg 26640
 cgtaattggc cccggccctt ggtgtaccag gaaattcccc agcccccacgac cgtactactt 26700
 ccccgagacg cccagggccga agtccagctg actaactcag gtgtccagct ggccggcgcc 26760

gcccgcctgt gtcgtcaccg ccccgcctag ggtataaaagc ggctggatcgatccaggcaga 26820
ggcacacagc tcaacgcacga ggtggtgagc tcttcgctgg gtctgcgacc tgacggagtc 26880
ttccaactcg ccggatcggg gagatcttcc ttacgcctc gtcaaggccgt cctgactttg 26940
gagagtttgt ctctcgacgc ccgcgtcggtt ggcacatcgca ctctccagtt cgtggaggag 27000
ttcactccct cggctctactt caacccttc tccggctccc ccggccacta cccggacgag 27060
ttcatcccga acttcgacgc catcagcgag tcgggtggacg gctacgattt aatgtccccat 27120
ggtggcgcag ctgaccttagc tcggcttca cacctggacc actgcccgg cttccgctgc 27180
ttcgctcggtt atctcgccga gtttgcctac tttagctgc ccgaggagca ccctcaggc 27240
ccggccacg gagtgcggat catcgctgaa gggggcctcg actcccacct gcttcggatc 27300
ttcagccagc gaccgatccct ggtcgagcgc gagcaaggac agacccttct gaccctgtac 27360
tgcatctgca accaccccg cctgcatgaa agtctttgtt gtctgctgtg tactgagtt 27420
aataaaaagct gagatcagcg actactccgg actcgattgtt ggtgttccctg ctatcaaccg 27480
gtccctgttc ttcaccggga acgagaccga gctccagttt cagtgttaagc cccacaagaa 27540
gtacccctacc tggctgttcc agggctcccc gatcggcggtt gtcaaccact ggcacaacga 27600
cgaggatccctg ctgagcggcc ccgccaacact tacttttcc acccgacagaa gcaagctcca 27660
gctcttccaa cccttcctcc ccgggaccta tcagtgcgtc tcgggacccct gccatcacac 27720
cttccacactg atcccgaata ccacagcgcc gctcccccgtt actaacaacc aaactaccca 27780
ccatcgccac cgtcgcgacc ttctctgaatc taacactacc acccacacccg gaggtgagct 27840
ccgaggtcga ccaacctctg ggattttacta cggccctgg gaggtggatgg ggttaatagc 27900
gcttaggccta gttgtgggtg ggcttttggc tctctgtac ctataccctc cttgtgttc 27960
gtacttagtg gtgtgtgtt gctggttaa gaatggggaa agatcacccct agttagctgc 28020
ggtgcgctgg tggcggtggg ggtgtttcg attgtggac tggcgccgc ggctgttagtg 28080
aaggagaagg ccgatccctg cttgcatttc aatcccgaca attgccagct gagtttcag 28140
cccgatggca atcggtgcgc ggtgctgatc aagtgcggat gggaaatgcga gaacgtgaga 28200
atcgagtaca ataacaagac tcggaaacaat actctcgctt ccgtgtggca gccccgggac 28260
cccgagtggtt acaccgtctc tgtcccccgtt gctgacggct ccccgccac cgtgaacaat 28320
actttcattt ttgcgcacat gtgcgcacacg gtcatgtggta tgagcaagca gtacgatatg 28380
tggcccccca cgaaggagaa catcggtggc ttctccatcg cttacagcgc gtgcacggcg 28440
ctaattcaccg ctatcgatgtg cctgagcattt cacatgctca tcgttatcc ccccaaaaaat 28500

aatgccgaaa aagagaaaaca gccataaacac gtttttcac acacccctttt cagaccatgg 28560
cctctgttaa attttgctt ttatggcca gtctcattac tgttataagt aatgagaaac 28620
tcactattta cattggcaact aaccacactt tagacggaat tccaaaatcc tcatggatt 28680
gctatttga tcaagatcca gacttaacta tagaactgtg tggtaacaag ggaaaaaaa 28740
caaggcattca ttaattaac ttaattgcg gagacaattt gaaattaatt aatatcacta 28800
aagagtatgg aggtatgtat tactatgtt cagaaaataa caacatgcag ttttatgaag 28860
ttactgtAAC taatcccacc acaccttagaa caacaacaac caccaccaca aaaactacac 28920
ctgttaccac tatgcagctc actaccaata acattttgc catgcgtcaa atggtaaca 28980
atagcactca acccaccacca cccagtgagg aaattcccaa atccatgatt ggcattattg 29040
ttgctgttgt ggtgtgcattt ttgatcatcg cttgtgcattt ggtgtactat gccttctgt 29100
acagaaagca cagactgaac gacaagctgg aacacttaact aagtgttcaa ttttatttt 29160
ttagaaccat gaagatccta ggccttttaa tttttctat cattacctct gctctatgca 29220
attctgacaa tgaggacgtt actgtcggt tcggaaccaa ttatacactg aaaggccag 29280
cgaagggtat gtttcgtgg tattgtggc ttggactga cgagcaacag acagagctct 29340
gcaatgctca aaaaggccaa acctcaattt cttaaatctc taattatcaa tgcaatggca 29400
ctgacttagt actgctcaat gtcacgaaag catatgctgg cagctacacc tgccctggag 29460
atgatactga gaacatgatt ttttacaaag tggaaagtggc tgatcccact actccaccc 29520
caccaccac aactactcac accacacaca cagaacaaac cacagcagag gaggcagcaa 29580
agttagccctt gcaggtccaa gacagttcat ttgttggcat taccctataa cctgatcagc 29640
ggtgtccggg gctgctcgat agccgcattt tcgggtgtc ttgggattt gcagtcataa 29700
tcatctgcattt gttcattttt gcttgcgtt atagaaggct ttaccgacaa aaatcagacc 29760
caactgctgaa cctctatgtt taatttttc cagagccatg aaggcagttt gcactctgtt 29820
ttttgttctt ttgattggca ctgttttag tgtagctttt ttgaaacaaa tcaatgttac 29880
tgagggggaa aatgtgacac tggtaggcgt agagggtgtc caaaatacca cctggacaaa 29940
atccatcta gatgggtgga aagaaatttgc cacctggaaat gtcagtactt atacatgtga 30000
aggagttaat cttaccatttgc tcaatgtcag ccaaatttcaaa aagggttggaa ttaaaggcc 30060
atctgttgtt gtttagcaata gtgggtacta tacccagcat actcttatct atgacattat 30120
agttataccatg cttgcctacac ctagccccacc tagcactacc acacagacaa cccacactac 30180

acaaacaacc acatacagta catcaaatca gcctaccacc actacaacag cagaggttgc 30240
 cagctcgctc ggggtccgag tggcattttt gatgttggcc ccatctagca gtcccaactgc 30300
 tagtaccaat gagcagacta ctgaattttt gtccactgtc gagagccaca ccacagctac 30360
 ctcgagtgcc ttctctaga ca cgc aatct atcctcgctt tcctctacac caatcagtcc 30420
 cgctactact cctacccccc ctattctccc cactcccctg aagcaaacag acggcgacat 30480
 gcaatggcag atcaccctgc tcattgtat cgggttggc atcctggccg tggtgtcta 30540
 ctacatcttc tgccgcccga ttccaaacgc gcacccgaag ccggcctaca agcccatcgt 30600
 tgtcgggcag ccggagccgc ttcaggtgga aggggtcta aggaatcttc tcttctctt 30660
 tacagtatgg tgattgaatt atgattccta gacaaatctt gatcactatt cttatctgcc 30720
 tcctccaagt ctgtgccacc ctcgcctcgg tgccaaacgc eagtccagac tgtattggc 30780
 ctttcgcctc ctacgtgcgc tttgcctca tcacctgcatt ctgctgtgt agcatagtct 30840
 gcctgcttat caccttcttc cagttcatttgc actggatctt tggtgcgcate gcctacctgc 30900
 gccaccaccc ccagtaccgc gaccagcggag tggcgcggct gtcaggatc ctctgataag 30960
 catgcgggct ctgctacttc tcgcgttctt gctgttagtg ctccccctgc ccgtcgaccc 31020
 ccggacccccc acccagtccc ccgaggaggt ccgcaaatgc aaattccaag aaccctggaa 31080
 attcctcaaa tgctaccgcc aaaaatcaga catgcattccc agctggatca tgatcattgg 31140
 gatcgtaac attctggctt gcaccctcat ctcccttgcatttaccctt gcttgactt 31200
 tggttgaaac tcgcccagagg cgctctatct cccgcctgaa cctgacacac caccacagca 31260
 acctcaggca cacgcactac caccaccacc acagcctagg ccacaataca tgcccatatt 31320
 agactatgag gccgagccac agcgacccat gctccccgtt attagttact tcaatctaac 31380
 cggcggagat gactgaccca ctggccaaca acaacgtcaa cgaccccttc ctggacatgg 31440
 acggccgcgc ctcggagcag cgactcgccc aacttcgcatt tcgcccagcag caggagagag 31500
 ccgtcaagga gctgcaggac ggcatacgcca tccaccagtg caagaaaggc atcttcgtcc 31560
 tggtgaaaca ggccaagatc tcctacgagg tcacccagac cgaccatcgc ctctcctacg 31620
 agctcctgca gcagcgcac aagttcacct gcctggcgg agtcaaccccc atcgatca 31680
 cccagcgtc gggcgataacc aagggttgca tccactgtcc ctgcgactcc cccgactgct 31740
 tccacactct gatcaagacc ctctgcggcc tccgcgacccct cctccccatg aactaatcac 31800
 ccccttatcc agtgaatataa agatcatatt gatgatttgat gttataaaaa aataaagaat 31860
 cacttacttg aaatctgata ccaggtctct gtcctatgttt tctgccaaca ccacttcact 31920

cccccttcc cagctctggc actgcaggcc ccggcggtc gcaaacttcc tccacaccct 31980
gaaggggatg tcaaattccct cctgtccctc aatcttcatt ttatcttcta tcagatgtcc 32040
aaaaagcgcg tccgggtgga tgatgacttc gaccccgctc acccctacga tgcagacaac 32100
gcaccgaccg tgcccttcat caacccccc ttctgtcttt cagatggatt ccaagagaag 32160
ccccctgggg tgctgtccct gcgtctggcc gatcccgta ccaccaagaa cggggaaatc 32220
accctcaagc tgggagatgg ggtggacctc gactcctcgg gaaaactcat ctccaacacg 32280
gccaccaagg cccggcccccc tctcagtttt tccaaacaaca ccatttcctt taacatggat 32340
acccctttt acaacaacaa tggaaagtta ggcatgaaag tcactgtcc actgaagata 32400
ctagacacag acttgctaaa aacacttgtt gtatgttatg gacaagggtt aggaacaaac 32460
accactggtg cccttggc ccaactagca tccccacttg cttttgatag caatagcaaa 32520
attgccctta atttaggcaa tggaccattg aaagtggatg caaatagact gaacatcaat 32580
tgcaatagag gactctatgt tactaccaca aaagatgcac tggaaagccaa tataagttgg 32640
gctaatgcta tgacatttat aggaaatgcc atgggtgtca atattgatac acaaaaaggc 32700
ttgcaatttg gcaccactag taccgtcgca gatgtaaaa acgcttaccc catacaaatc 32760
aaacttggag ctggtctcac atttgacagc acaggtgcaa ttgttgcatg gaacaaagat 32820
gatgacaagc ttacactatg gaccacagcc gacccctctc caaattgtca catatattct 32880
gaaaaggatg ctaagcttac actttgttg acaaagtgtg gcagtcagat tctggcact 32940
gtttccctca tagctgttga tactggcagt ttaatccca taacaggaac agtaaccact 33000
gctctgtct cacttaaattt cgatgcaaat ggagtttgc aaagcagctc aacactagac 33060
tcagactatt ggaatttcag acagggagat gttacacctg ctgaagccta tactaatgct 33120
atagtttca tgcccaatct aaaagcatac cctaaaaaca caagtggagc tgcaaaaagt 33180
cacattgttgc gaaaaagtgttgc cctacatggg gatacaggca aaccactggc cctcattatt 33240
actttcaatg aaacaagtga tgaatcttgc acttactgtt ttaactttca atggcagtgg 33300
ggggctgatc aatataaaaaa tgaaacactt ggcgtcagtt cattcacctt ttccttatatt 33360
gctaaagaat aaaccccaact ctgtacccca tctctgtcta tggaaaaaac tctgaaacac 33420
aaaataaaaat aaagttcaag tggtttatttgc attcaacagt tttacaggat tcgagcagtt 33480
attttccttc caccctccca ggacatggaa tacaccaccc tctccccccg cacagccttg 33540
aacatctgaa tgccattggt gatggacatg cttttgtct ccacgttcca cacagttca 33600

gagcgagcca gtctcgggtc ggtcaggag atgaaaccct ccgggcactc ccgeatctgc 33660
acctcacagc tcaacagctg aggattgtcc tgggtggtcg ggatcacggt tatctggaag 33720
aagcagaaga gcggcggtgg gaatcatagt ccgcgaacgg gatcgccgg tggtgtcgca 33780
tcagggcccg cagcagtcgc tgtcgcccgc gctccgtcaa gctgctgctc agggggtccg 33840
ggtccaggga ctcccccagc atgatgccca cggccctcag catcagtcgt ctggtgccgc 33900
gggcgcagca gcgcatgcgg atctcgctca ggtcgtgca gtacgtgcaa cacaggacca 33960
ccaggttgtt caacagtcaca tagttcaaca cgctccagcc gaaactcatc gcgggaagga 34020
tgctacccac gtggccgtcg taccagatcc tcaggtaaat caagtggcgc cccctccaga 34080
acacgctgcc catgtacatg atctcccttgg gcatgtggcg gttcaccacc tcccggtacc 34140
acatcaccct ctgggtgaac atgcagcccc ggatgatcgt gcggAACAC agggccagca 34200
ccgcggccgc cgccatgcag cgaagagacc cgggtcccg acaatggcaa tggaggaccc 34260
accgctcgta cccgtggatc atctggagc tgaacaagtc tatgttgca cagcacaggc 34320
atatgtcat gcatctcttc agcactctca gtcctcggg ggtcaaaacc atatcccagg 34380
gcacggggaa ctcttgcagg acagcgaacc cgcagaaca gggcaatcct cgacataac 34440
ttacattgtg catggacagg gtatcgcaat cagggcagcac cgggtgatcc tccaccagag 34500
aagcgcgggt ctcggcttcc tCACAGCGTG gtaagggggc cggccgatac ggggtatggc 34560
gggacgcggc tgatcggtt cgccaccgtg ttatgtgca gttgtttcg gacatttcg 34620
tacttgctgt agcagaacct ggtccggcg ctgcacaccg atgcggcgcg ggggtcccg 34680
cgcttggAAC gtcgggtgtt gaagtgtaa aacagccact ctctcagacc gtgcagcaga 34740
tcttagggct caggagtgtat gaagatccca tcatgcctga tggctctaatac acatcgacc 34800
accgtggaat gggccagacc cagccagatg atcaatttt gttgggttcc ggtgacggcg 34860
ggggagggaa gaacaggaag aaccatgatt aacttttaat ccaaacggtc tcggagact 34920
tcaaaatgaa gatcgccggag atggcacctc tcggcccccgc tgggttggtg gaaaataaca 34980
gccaggtcaa aggtgatacg gttctcgaga tggccacgg tggcttccag caaaggctcc 35040
acgcgcacat ccagaaacaa gacaatagcg aaagcggag gttctctaa ttccctcaatc 35100
atcatgttac actccctgcac catccccaga taattttcat tttccagcc ttgaatgatt 35160
cgaactagtt cctgaggtaa atccaaagcca gccatgataa agagctcgcg cagagcggcc 35220
tccacccggca ttcttaagca caccctcata attccaagat attctgctcc tggttcaccc 35280
gcagcagatt gacaagcgga atatcaaaat ctctgcggcg atccctaagc tcctccctca 35340

gcaataactg taagtactct ttcatatcct ctccgaaatt tttagccata ggaccaccag 35400
 gaataagatt agggcaagcc acagtacaga taaaccgaag tcctccccag tgagcattgc 35460
 caaatgcaag actgtataa gcatgctggc tagaccgggt gatatcttcc agataactgg 35520
 acagaaaatc gcccaggcaa ttttaagaa aatcaacaaa agaaaaatcc tccaggtgca 35580
 cgtttagagc ctcgggaca acgatggagt aaatgcaagc ggtgcgttcc agcatggta 35640
 gttagctgat ctgtagaaaa aaacaaaaat gaacattaaa ccatgcttagc ctggcgaaca 35700
 ggtgggtaaa tcgttcttc cagcaccagg caggccacgg ggtctccggc acgaccctcg 35760
 taaaaattgt cgctatgatt gaaaaccatc acagagagac gttccggtg gccggcgtga 35820
 atgattcgac aagatgaata caccccccga acattggcgt ccgcgagtga aaaaaagcgc 35880
 ccaaggaagc aataaggcac tacaatgctc agtctcaagt ccagcaaagc gatgccatgc 35940
 ggatgaagca caaaattctc aggtgcgtac aaaatgtaat tactccctc ctgcacaggc 36000
 agcaaagccc ccgatccctc caggtacaca tacaagcct cagcgtccat agcttaccga 36060
 gcagcagcac acaacaggcg caagagtcag agaaaggctg agctctaacc tgtccacccg 36120
 ctctctgctc aatatatagc ccagatctac actgacgtaa aggccaaagt ctaaaaatac 36180
 ccgccaata atcacacacg cccagcacac gcccagaaac cggtgacaca ctcaaaaaaaaa 36240
 tacgcgcact tcctcaaacg cccaaactgc cgtcatttcc gggttccac gctacgtcat 36300
 caaaaattcga ctttcaaatt ccgtcgaccg ttaaaaacgt cggccggccc gcccttaacg 36360
 gtcgcgcctc ccgcagccaa tcaccggccc gcattcccaa attcaaatac ctcatttgca 36420
 tattaacgcg cacaaaaagt ttgaggtata ttattgatga tg 36462

<210> 2
 <211> 36604
 <212> DNA
 <213> chimpanzee adenovirus serotype Pan6

<400> 2
 catcatcaat aatataccctc aaacttttgg tgcgcgtaa tatgcaaatg agctgtttga 60
 atttggggag ggaggaaggt gattggctgc gggagcggcg accgttaggg gcggggcggg 120
 tgacgttttgc atgacgtggc tatgaggcgg agccggtttgc caagttctcg tggaaaaagt 180
 gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa tttcccgcg ctctctgaca 240
 ggaaatgagg tggattctggg cggatgcaag tgaaaacggg ccatttcgc gcgaaaactg 300
 aatgaggaag tgaaaatctg agtaatttcg cgtttatggc agggaggagt attgcccag 360

ggccgagtag actttgaccg attacgtggg ggtttcgatt accgtatTTT tcacctaAT	420
ttccgcgtac ggtgtcaaAG tccgggtttt ttacgttagc gtcagctgat cgccaggGta	480
ttaaacCTG cgctctCTAG tcaagaggCC actcttgagt gccagcgagt agagTTTct	540
cctccgcGCC gcgagtcaga tctacacttt gaaagatgag gcacctgaga gacctgcccG	600
gtaatgtttt cctggctact gggAACgaga ttctggatt ggtgggtggac gccatgatGG	660
gtgacgacCC tccagagccc CCTACCCat ttgaggcGCC ttcgctgtac gatttGTatG	720
atctggaggt ggatgtGCC gagagcgacc ctaacgagga ggCGGTGAAT gatttGTTA	780
gcgtGCCGC gctgtggct GCCGAGCAGG ctaatacggA ctctggctCA gacagcgatt	840
cctctctCCA tacCCCgaga CCCGGAGAG gtgagaaaaa gatccccGAG cttaaaggGG	900
aagagctcga CCTGCGCTGC tatgaggaat gcttcctcc gagcgtatGat gaggaggacG	960
aggaggcGat tcgagctgcg gtgaaccagg gagtgaaaac tgcccggcAG agctttAGCC	1020
tggactgtcc tactctGCC ggacacggct gtaagtcttG tgaatttcat CGCATGAATA	1080
ctggagataa gaatgtgatG tGtgcctgt gctatatGAG agcttacaac cattgtgttt	1140
acagtaagtG tgattaACTT tagttggaa ggcaGAGGGT gactgggtgc tgactggTTT	1200
atttatgtat atgtttttt atgtgttagt cccgtctctG acgttagatGA gacccccact	1260
tcagagtGca tttcatcacc CCCAGAAATT ggcgagGAAC CGCCCGAAGA tattattcat	1320
agaccAGttG cagtGAGAGt CACCGGGCGG agAGCAGCTG tggagatTT ggatGacttG	1380
ctacaggGTG gggatGAACC tttggacttG tGtacCCGGa aacGCCCGAG GCACTAAGtG	1440
ccacacatgt gtgtttactt aaggGtGatGt cagtatttAt agggtgtGGA gtGcaataAA	1500
atccGtGttG actttAAGtG cgtgtttat gactcaggGG tggggactGt gggtatataA	1560
gcaggtGcAG acctgtgtGG tcagttcAGA gcaggactCA tggagatctG gactgtcttG	1620
gaagactttc accagactAG acagtGctA gagaactcat CGGAGGGAGt CTCTTACCTG	1680
tggagattct gtttcggGtG gcctctAGtC aagctAGtCt ataggGcAA acaggattat	1740
aaggAACAt ttGAGGAtAt tttgagAGAG tGtccGgtA tttttGactC tctcaacttG	1800
ggccatcAGt CTCACttaA ccAGAGtAtt ctGAGAGCCC ttGactttC tactctGgc	1860
agaactACG CCGCGGtAGC CTTTTGcc tttattcttG acAAatGGAG tcaAGAAACC	1920
catttcAGCA gggattACG tctggactGc ttAGCAGTAGt CTTGtGGAG AACATGGAGG	1980
tGCCAGCGCC tGAATGCAAT CTCGCGtAC ttGCCAGtAC AGCCGGTAGA cacGtGAGG	2040

atcctgagtc	tccagtcaacc	ccaggaacac	caacggccgc	agcagccgca	gcaggagcag	2100
cagcaagagg	aggaccgaga	agagaacccg	agagccggtc	tggaccctcc	ggggcggag	2160
gaggaggagt	agctgacttg	tttcccggc	tgccgggt	gctgacttagg	tctccagtg	2220
gacggggag	ggggatataag	cgggagaggc	atgaggagac	tagccacaga	.actgaactga	2280
ctgtcagtct	gatgagccgc	aggccccag	aatcggtgt	gtggcatgag	gtcagtcgc	2340
aggggataga	tgaggtctcg	gtgatgcatt	agaaatattc	cctagaacaa	gtcaagactt	2400
gttgggttgg	gcccgaggat	gattgggagg	tagccatca	gaattatgcc	aagctggctc	2460
tgaagccaga	caagaagtac	aagattacca	aactgattaa	tatcagaaat	tcctgctaca	2520
tttcaggaa	tggggcccgag	gtggagatca	gtacccagga	gagggtggcc	ttcagatgtt	2580
gtatgatgaa	tatgtacccg	ggggtgttgg	gcatggaggg	agtcacccctt	atgaacacga	2640
ggttcagggg	tgtgggtat	aatggggtgg	tctttatggc	caacaccaag	ctgacagtgc	2700
acggatgctc	cttctttggc	ttcaataaca	tgtgcattca	ggcctgggc	agtgtttcag	2760
tgaggggatg	cagctttca	gccaaactgga	tgggggtcgt	gggcagaacc	aagagcaagg	2820
tgtcagtgaa	gaaatgcctg	tgcagaggt	gccacctggg	ggtgatgagc	gagggcgaag	2880
ccaaagtcaa	acactgcgcc	tctaccgaga	cgggctgtt	tgtgcattca	aaggcaatg	2940
cccaagtcaa	gcataacatg	atctgtgggg	cctcggatga	gcgcggctac	cagatgctga	3000
cctgcgcgg	tggaaacagc	catatgctgg	ccacccgtca	tgtggcctcg	caccccgca	3060
agacatggcc	cgagttcgag	cacaacgtca	tgacccgtc	caatgtgcac	ctgggctccc	3120
gccgaggcat	gttcatgcc	taccagtca	acatgcatt	tgtgaaggtg	ctgctggagc	3180
ccgatgccat	gtccagagt	agcctgacgg	gggtgttga	catgaatgt	gagctgttgg	3240
aaattctgag	atatgatgaa	tccaagacca	ggtgccgggc	ctgcgaatgc	ggaggcaagc	3300
acgccaggct	tcaagccgt	tgtgtggagg	tgacggagga	cctgcgaccc	gatcatttgg	3360
tgttgtccctg	caacgggacg	gagttcggt	ccagcgggga	agaatctgac	tagatgtgt	3420
agtgtttggg	gctgggtgt	agcctgcatg	aggggcagaa	tgactaaaat	ctgtggttt	3480
ctgtgtgttgc	cagcagcatg	agcggaaagcg	cctccttga	gggaggggtt	ttcagccctt	3540
atctgacggg	gcgtctcccc	tctggccgg	gagtgcgtca	gaatgtgtat	ggatccacgg	3600
tggacggccg	gcccgtgcag	cccgcgaact	cttcaaccct	gacctacgcg	accctgagct	3660
cctcgtccgt	ggacgcagct	gcccgcgcag	ctgctgttcc	cgccgcacgc	gccgtgcgcg	3720
gaatggccct	gggcgcggc	tactacagct	ctctggtggc	caactcgagt	tccaccaata	3780

atcccggccag	cctgaacgag	gagaagctgc	tgctgtgtat	ggcccagtc	gaggccctga	3840
cccagcgccct	gggcgagctg	accacgcagg	tggctcagct	gcaggcggag	acgcggggccg	3900
cggttgccac	ggtgaaaacc	aaataaaaaa	tgaatcaata	aataaacgga	gacggttgtt	3960
gattttaca	cagagtcttg	aatctttatt	tgattttcg	cgcgcggtag	gccctggacc	4020
acgggtctcg	atcattgagc	acccggtgga	tctttccag	gaccggtag	agggtggctt	4080
ggatgttgag	gtacatggc	atgagccgt	cccgggggtg	gaggtagctc	cattgcaggg	4140
cctcgtgttc	ggggatggtg	ttgttaatca	cccaagtata	gcaggggcgc	agggcgtggt	4200
gctgcacgt	gtccttgagg	aggagactga	tggccacggg	cagccccctg	gtgttaggtt	4260
tgacgaacct	gtttagctgg	gaggatgca	tgcgggggg	gatgagatgc	atctggcct	4320
ggatcttgag	attggcgatg	ttcccgccca	gatcccgcg	ggggttcatg	ttgtgcagga	4380
ccaccagcac	ggtgtatccg	gtgcacttgg	ggaatttgc	atgcaacttgc	gaagggaaagg	4440
cgtgaaagaa	tttggagacg	cccttgtac	cgcccaggtt	ttccatgcac	tcatccatga	4500
tgatggcgat	gggcccgtgg	gcggcggcct	ggcaaagac	gtttcggggg	tccggacacat	4560
cgtagttgtg	gtcctgggtg	agctcgat	aggccatttt	aatgaatttgc	gggcggaggg	4620
tgcccgaactg	ggggacgaag	gtgcctcga	tccggggggc	gtagttgccc	tgcagatct	4680
gcatctccca	ggccttgagc	tccgggggg	ggatcatgtc	cacctgcggg	gcgtgaaaa	4740
aaacggtttc	cggggcgggg	gagatgagct	gggcccggaaag	caggttccgg	agcagctggg	4800
acttgcgcga	accgggtgggg	ccgtagatga	ccccgatgac	cggctgcagg	tggttagttga	4860
gggagagaca	gctgcccgtcc	tccggggagga	ggggggccac	ctcggttcatc	atctcgcgca	4920
catgcatgtt	ctcgcgcaacg	agttccgcca	ggaggcgctc	gcggggcagc	gagaggagct	4980
cttgcagcga	ggcgaagttt	ttcagccgct	tgagtccgtc	ggccatgggc	attttggaga	5040
gggtctgttg	caagagttcc	agacggtccc	agagctcggt	gatgtgtct	agggcatetc	5100
gatccagcag	acctctctgt	ttcgcgggtt	ggggcgactg	cgggagtagg	gcaccaggcg	5160
atggcggtcc	agcgaggcca	gggtccggtc	cttccagggc	cgcagggtcc	gcgtcagcgt	5220
ggtctccgtc	acggtaagg	ggtgcgccgc	gggtggggcg	cttgcgaggg	tgcgtttcag	5280
gctcatccgg	ctggcgaga	accgctcccg	gtcgccgc	tgcgcgtcgg	ccaggtagca	5340
attgagcatg	agttcgtagt	tgagcgccctc	ggccgcgtgg	cccttggcgc	ggagcttacc	5400
tttggaaagt	tgtccgcaga	cgggacagag	gagggacttg	agggcgtaga	gcttgggggc	5460

gaggaagacg	gactcgaaaa	cgttaggcgtc	cgccgcgcag	ctggcgcaaa	cggtctcgca	5520
ctccacgac	caggtgagg	cggggcggtt	ggggtaaaaa	acgaggtttc	ctccgtgctt	5580
tttgcgtt	ttcttaccc	tggctccat	gagctcggt	ccccgctggg	tgacaaagag	5640
gctgtccgt	tccccgtaga	ccgactttat	gggccggtcc	tcgagcgaaa	tgccgcggtc	5700
ctcgtag	aggaaccccg	cccactccga	gacgaaggcc	cggtccagg	ccagcacgaa	5760
ggaggccacg	tgggaggggt	agcggtcgtt	gtccaccage	gggtccacct	tctccagggt	5820
atgcaagcac	atgtccccct	cgtccacatc	caggaaggtg	attggcttgt	aagtgttaggc	5880
cacgtgaccg	ggggtcccgg	ccgggggggt	ataaaagggg	gcggggccct	gtcgctcctc	5940
actgtcttcc	ggatcgctgt	ccaggagcgc	cagctgttgg	ggttaggtatt	ccctctcgaa	6000
ggcgggcatg	acctcggcac	tcaggttgtc	agtttctaga	aacgaggagg	atttgcattt	6060
gacggtgcgc	ttggagacgc	ctttcatgag	cccctcgcc	atttggtcag	aaaagacgt	6120
cttttgcgtt	tcgagcttgg	tggcgaagga	gccgttaggg	gcgttggaga	gcagcttggc	6180
gatggagcgc	atggtctgg	tctttcctt	gtcggcgcgc	tccttggcgg	cgatgtttag	6240
ctgcacgtac	tcgcgcgcac	cgcacttcca	ttcggggaaag	acgggtgtga	gtcgctcgaa	6300
cacgattctg	acccggccagc	cgcgggtgt	cagggtgtatg	aggccacgc	tggtggccac	6360
ctcgccgcgc	aggggctcg	tggccagca	gaggcgcccg	cccttgcgcg	agcagaagggg	6420
ggcagcgaaa	tccagcatga	gtcgtcgaaa	ggggtcggcg	tccacgggtga	agatgcgggg	6480
caggagctcg	gggtcgaagt	agctgtatc	ggtgtccaca	ttgttccagcg	ccgcttgcac	6540
gtcgccgcac	gccagcgcgc	gtcgtaggg	gtgaggggc	gtgcccagg	gcatgggggt	6600
cgtgagcgcgc	gaggcgtaca	tgccgcagat	gtcgtagacg	tagagggtgt	cctcgaggac	6660
gccgatgtat	gtggggtagc	agcgcccccc	gccccatgt	gcgcgcacgt	agtcgtacag	6720
ctcggtcgag	ggcgcgagga	gccccgtgcc	gagggtggag	cggtgcggct	tttcggcgcgc	6780
gtagacgatc	tggcgaaaga	tggcggtgg	gttggaggag	atgggtggcc	tttggaaagat	6840
gttgaagtgg	gcgtggggca	ggccgaccga	gtccctgtatg	aagtgggcgt	aggagtcctg	6900
cagttggcg	acgagctcg	cggtgacgag	gacgtccagg	gcgcagtatg	cgagggtctc	6960
ttggatgtat	tcataacttga	gtggccctt	ctcgatccac	agctcgccgt	tgagaaggaa	7020
ctcttcgcgg	tccttccagt	actcttcgag	ggggaaacccg	tcctgtatcg	cacggtaaga	7080
gcccaccatg	tagaactgg	tgacggcctt	gtaggcgcag	cagcccttct	ccacggggag	7140
ggcgtaagct	tgcgccgcct	tgcgccaggaa	ggtgtgggtg	agggcgaagg	tgtcgccac	7200

catgaccttg	aggaactggt	gcttgaagtc	gaggtcgtcg	cagccgcctt	gctcccagag	7260	
tttggaaagtcc	gtgcgccttct	tgttaggcggg	gttaggc当地	gcttgc当地	7320		
gaggatcttg	ccccgc当地	gcatgaagtt	gc当地gtatgc	cgaaaaggct	ggggcacctc	7380	
ggcccggttg	ttgatgc当地	gggc当地ggc当地	gacgatctcg	tc当地agc当地	tgatgttg	7440	
cccgacgatg	tagagttcca	c当地aatgc当地	gc当地ggc当地	acgtgggca	gcttcttgag	7500	
ctcgctgtag	gtgagctcg	cggggtcg	gagcccg	tgctcgaggg	ccc当地gtcg	7560	
gacgtggggg	ttggcgctga	ggaaggaa	act	ccagagatcc	acggccaggg	cggtctgcaa	7620
gc当地gtcccg	tactgacg	actgttgcc	cacggccatt	tttgc当地ggg	tgacgc当地	7680	
gaaggtgc当地	gggtcgcc	gccagcg	ccacttg	tggagggc当地	ggtcg	7740	
gagctcgacg	acgccc当地	ccccggag	tttcatgacc	agcatgaagg	ggacgagctg	7800	
cttgc当地gaag	gacccc当地	agg	ttccacatcg	taggtgagga	agaggc当地	7860	
ggtgc当地gagga	tgcgagcc	tg	ggatctcc	tgccacc	tgaggaaatg	7920	
gctgttgc当地	tgatg	gaaatgc当地	acggccg	gagcactcg	gcttgc当地	7980	
atacaaggc当地	cccg	actgt	cgcaacgc当地	cacggat	acgtg	8040	
ctgggttcc	ttggcgagga	at	ttcagtg	gc当地gtgg	gctggcg	gcatctcg	8100
ctgtactacg	tcttgc当地	c当地ggc当地	atcg	tctg	tc当地gt	8160	
gagccc当地	gggaggc当地	tccagac	cttgc当地	ggctcg	ggagag	cgaggac	8220
ggc当地	ccggagctg	ccagg	gtc当地	ggatc	c当地gtgg	8280	
cgccggc当地	cggtt	gactt	gc当地	ttccagg	cgccgg	ccagatgg	8340
cttgc当地	acggc当地	ccgt	ttgtgg	gtccac	tc当地	8400	
ggc当地	accgtg	cccc	gttctt	ggcg	cttgc当地	tc当地	8460
cgccgaggac	g	ggc当地	gggg	ggctcg	ggggc当地	8520	
ggc当地	ggc当地	gggg	ggatc	ggggc当地	ggggc当地	8580	
g	ggc当地	gggg	tc当地	ggatc	ggggc当地	8640	
gtgagttga	ac	ctgaaaga	gat	tc当地	ggatc	8700	
tgccgc当地	tcttgc当地	gtcgccc	tgat	tc当地	ggatc	8760	
tgctcgatct	cctc	ctc	tc当地	ggatc	ggatc	8820	
tc当地	ggg	ggatc	ggatc	ggatc	ggatc	8880	

cggctgtaga	ccacggctcc	gtcggggtcg	cgcgcgcga	tgaccacctg	ggcgaggttg	8940
agctcgacgt	ggcgctgaa	gaccgcgtag	ttgcagaggc	gctggtagag	gtatgtgagc	9000
gtggtggcga	tgtgctcggt	gacgaagaag	tacatgatcc	agcggcggag	cgccatctcg	9060
ctgacgtcgc	ccagggcttc	caagcgttcc	atggcctcg	agaagtccac	ggcgaagttg	9120
aaaaactggg	agttgcgcgc	cgagacggc	aacteectcct	ccagaagacg	gatgagctcg	9180
gcgatggtgg	cgcgcaccc	gcgctcgaag	gccccggggg	gctcccttc	catcteetcc	9240
tcttcctcct	ccactaacat	ctcttctact	tcctcctcag	gaggcggtgg	cgggggaggg	9300
gcctcgctc	gcggcgccc	cacggcaga	cggtcgatga	agcgctcgat	ggtctccccg	9360
cgcggcgac	gcatggtctc	ggtgacggcg	cgcccgctct	cgcgggccg	cagcatgaag	9420
acgcgcgcgc	gcatctccag	gtggccgccc	ggggggctc	cgttggcag	ggagagggcg	9480
ctgacgatgc	atcttatcaa	ttgaccctgta	gggactccgc	gcaaggacct	gagcgtctcg	9540
agatccacgg	gatccgaaa	ccgctgaacg	aaggcttcga	gccagtcga	gtcgcaaggt	9600
aggctgagcc	cggtttctg	ttcttcgggt	atttggtcgg	gaggcggcg	ggcgatgctg	9660
ctggtgatga	agttgaagta	ggcggtcctg	agacggcgga	tggtggcgag	gagcaccagg	9720
tccttggcc	cggcttgctg	gatgcgcaga	cggtcgccca	tgccccaggc	tggttcctga	9780
cacctggcga	ggtccttcta	gtagtcctgc	atagcccgct	ccacgggcac	ctcctcctcg	9840
cccgcgccgc	cgtgcacgcg	cgtgagcccg	aaacccgcgt	gcggctggac	gagcgcagg	9900
tcggcgacga	cgcgcgtcggt	gaggatggcc	tgctggatct	gggtgagggt	ggtctggaag	9960
tcgtcgaagt	cgacgaagcg	gtggtaggct	ccgggtttga	tggtgttagga	gcagttggcc	10020
atgacggacc	agttgacggt	ctggtgccg	ggtcgcacga	gctcgtggta	cttgaggcgc	10080
gagtaggcgc	cggtgtcgaa	gatgtagtcg	ttgcaggcgc	gcacgaggta	ctggatccg	10140
acgaggaagt	cgggcgccgg	ctggcggtag	agcggccatc	gctcggtggc	gggggcgcgg	10200
ggcgcgaggt	cctcgagcat	gaggcggtgg	tagccgtaga	tgtacctgga	catccaggtg	10260
atgcccggcgg	cgggtggtgga	ggcgccgggg	aactcgccga	cgcggttcca	gatgttgcgc	10320
agcggcagga	agtagttcat	gtggccgcg	gtctggcccg	tgaggcgcgc	gcagtcgtgg	10380
atgctctaga	catacgggca	aaaacgaaag	cggtcagcgg	ctcgactccg	tggcctggag	10440
gctaagcgaa	cgggttgggc	tgcgcgtgta	ccccggttcg	aatctcgaaat	caggctggag	10500
ccgcagctaa	cgtggtactg	gcactcccgt	ctcgacccaa	gcctgctaac	gaaacctcca	10560
ggataacggag	cggggtcggt	ttttggcctt	ggtcgtgg	cataaaaaac	tagtaagcgc	10620

ggaaagcggc cgccccgcgtat ggctcgctgc cgtagtctgg agaaagaatc gccagggttg 10680
 cgttcggtg tgccccggtt cgagccctcag cgctcgccgc cggccggatt ccgcggctaa 10740
 cgtggcgtg gctgccccgt cgtttccaag acccccttagc cagccgactt ctccagttac 10800
 ggagcgagcc ccttttttt tttcttgtgt ttttgcaga tgcattccgt actgcggcag 10860
 atgcgcccccc accctccacc acaaccgccc ctaccgcagc agcagcaaca gccggcgctt 10920
 ctgccccccgc cccagcagca gccagccact accgcggcgg cccgcgtgag cggagccggc 10980
 gttcagtatg acctggcctt ggaagagggc gaggggctgg cgccgcgtgg ggcgtcgctg 11040
 ccggagcggc accccgcgcgt gcagatgaaa agggacgcctc gcgaggccta cgtgccaaag 11100
 cagaacctgt tcagagacag gagcggcggag gagcccgagg agatgcgcgc ctccgccttc 11160
 cacgcggggc gggagctgcg gcgcggcctg gaccgaaagc gggtgctgag ggacgaggat 11220
 ttcgaggcgg acgagctgac gggatcagc cccgcgcgcg cgacgtggc cgccggcaac 11280
 ctggcacgg cgtacgagca gaccgtgaag gaggagagca acttccaaaa atccttcaac 11340
 aaccacgtgc gcacgctgat cgccgcgcgag gaggtgaccc tgggcctgat gcacctgtgg 11400
 gacctgctgg aggccatcgt gcagaacccc acgagcaagc cgctgacggc gcagctgttt 11460
 ctggtggtgc agcacagtcg ggacaacgag acgttcaggg aggcgcgtct gaatatcacc 11520
 gagcccgagg gccgctggct cctggacctg gtgaacattt tgcaagcat cgtggtgcaag 11580
 gagcgcgggc tgccgcgtc cgagaagctg gcggccatca acttctcggt gctgagtctg 11640
 ggcaagtaact acgcttagaa gatctacaag accccgtacg tgcccataga caaggaggtg 11700
 aagatcgacg ggttttacat ggcacatgacc ctgaaagtgc tgaccctgag cgacgatctg 11760
 ggggtgtacc gcaacgcacag gatgcaccgc gcggtgagcg ccagccgcg ggcgcgactg 11820
 agcgaccagg agctgatgca cagcctgcag cggccctga ccggggccgg gaccgagggg 11880
 gagagctact ttgacatggg cgccgacctg cgctggcage ccagccgcg ggccttggaa 11940
 gctgcccggcg gttcccccata cgtggaggag gtggacgtat aggaggagga gggcgagtac 12000
 ctggaagact gatggcgca ccgtatTTT gctagatgca gcaacagcca ccgcgcgcgc 12060
 ctcctgatcc cgcgatgcgg gcggcgctgc agagccagcc gtccggcatt aactcctcg 12120
 acgattggac ccaggccatg caacgcatac tggcgctgac gacccgcaat cccgaagcct 12180
 tttagacagca gcctcaggcc aaccggctct cggccatctt ggaggccgtg gtgcctcgc 12240
 gctcgaaaccc cacgcacgag aaggtgctgg ccatcgtcaa cgcgctggg gagaacaagg 12300

ccatccgcgg	tgacgaggcc	gggctggtgt	acaacgcgt	gctggagcgc	gtggcccgt	12360
acaacacgcac	caacgtgcag	acgaacctgg	accgcattgt	gaccgacgtg	cgcgaggcgg	12420
tgtcgacgcg	cgagcggttc	caccgcgagt	cgaacctggg	ctccatggtg	gcgctgaacg	12480
ccttctcgag	cacgcagccc	gccaacgtgc	cccggggcca	ggaggactac	accaacttca	12540
tcaagcgcgt	gcccgtatg	gtggccgagg	tgccccagag	cgaggtgtac	cagtcggggc	12600
cggactactt	tttccagacc	agtcgccagg	gcttgcagac	cgtgaacctg	agccaggctt	12660
tcaagaactt	gcagggactg	tggggcgtgc	aggccccggt	cggggaccgc	gacgcgggt	12720
cgagcctgct	gacgcccgaac	tgcgcctgc	tgcgtgtgt	ggtggcgccc	ttcacggaca	12780
gcggcagcgt	gagccgcgac	tgcgtacctgg	gctacctgt	taacctgtac	cgcgaggcca	12840
tgcgcacaggc	gcacgtggac	gagcagaccc	accaggagat	cacccacgtg	agccgcgcgc	12900
tgggccagga	ggacccgggc	aacctggagg	ccaccctgaa	tttccctgt	accaaccgg	12960
cgcagaagat	cccgccccag	tacgcgtga	gcaccgagga	ggagcgcatac	ctgcgtacg	13020
tgcagcagag	cgtggggctg	tttccctgt	aggagggggc	cacgcggcage	gcggcgctcg	13080
acatgaccgc	gchgcaacatg	gagcccgac	tgtacgccc	caaccggcc	ttcatcaata	13140
agctgtatgga	ctacttgcac	cggggcccg	ccatgaactc	ggactactt	accaacgcca	13200
tcttgaaccc	gcactggctc	ccggccggcc	ggttctacac	gggcgagttac	gacatggcc	13260
accccaacga	cgggttcctg	tgggacgacg	tggacagcag	cgtgttctcg	ccgcgtccag	13320
gaaccaatgc	cgtgttggaa	aaagagggcg	gggacccggcg	gcccgttcc	gcgcgttcc	13380
gtcgcgccgg	tgcgtccgcg	gcgggtcccc	aggccgccc	cccttcccc	agcctgcct	13440
tttcgtgaa	cagcgtgcgc	agcagcgcac	tgggtcggct	gacgcgacc	cgccctgt	13500
gcgaggagga	gtacctgaac	gactcctt	tgaggccc	gcgcgagaag	aacttcccc	13560
ataacgggat	agagagcc	gtggacaaga	tgagccgt	gaagacgtac	gcgcacgag	13620
acagggacga	gccccgagct	agcagcgcag	gcacccgt	acgcacgg	cacgcacaggc	13680
agcggggact	ggtgtggac	gatgaggatt	ccgcccacg	cagcagcgt	ttggacttgg	13740
gtgggagtgg	tggtaaccc	tgcgttcc	tgccccc	tatcgccgc	ctgtatgt	13800
aatctgaaaa	aataaaagac	ggtactcacc	aaggccatgg	cgaccagcgt	gcgttcttct	13860
ctgttgttt	tagtagtatg	atgagggcg	tgtacccgg	gggtccct	ccctcgta	13920
agagcgtat	gcagcaggcg	gtggccggcg	cgatgcagcc	cccgctggag	gcgccttacg	13980
tgccccccgcg	gtacctggcg	cctacggagg	ggcggaaacag	cattcgta	tcggagctgg	14040

cacccttgtta cgataaccacc cggttgtacc tggtgacaa caagtcggca gacatcgcc 14100
 cgctgaacta ccagaacgac cacagcaact tcctgaccac cgtggtgcag aacaacgatt 14160
 tcaccccccac ggaggccagc acccagacca tcaacttga cgagcgctcg cgggtggggcg 14220
 gccagctgaa aaccatcatg cacaccaaca tgcccaacgt gaacgagttc atgtacagca 14280
 acaagttcaa ggcgcggtg atggtctcgc gcaagacccc caacgggtg gatgatgatt 14340
 atgatggtag tcaggacgag ctgacctacg agtgggtgga gtttgagctg cccgagggca 14400
 acttctcggt gaccatgacc atcgatctga tgaacaacgc catcatcgac aactacttgg 14460
 cgggtggggcg gcagaacggg gtgctggaga gcgacatcg cgtgaagttc gacacgcgca 14520
 acttccggct gggctgggac cccgtgaccc agctggtgat gccgggctg tacaccaacg 14580
 aggccttcca cccccacatc gtcctgctgc cccgctgccc cgtggacttc acggagagcc 14640
 gcctcagcaa cctgctgggc atccgcaagc ggcagccctt ccaggagggc ttccagatcc 14700
 tgtacgagga cctggagggg ggcaacatcc cccgcgttcc ggtatgtcgaa gcctacgaga 14760
 aaagcaagga ggatagcacc gcccggcga cccgagccgt ggccacccgc tctaccgagg 14820
 tgccggggcga taattttgtc agcgctgcgg cagcggccga ggcggctgaa accgaaagta 14880
 agatagtcat ccagccgggtg gagaaggaca gcaaggacag gagctacaac gtgtcgccg 14940
 acaagaaaaa caccgcctac cgcagctggt acctggctca caactacggc gaccccgaga 15000
 agggcgtgcg ctccctggacg ctgctcacca cctcggacgt cacctgcggc gtggagcaag 15060
 tctactggtc gtcgtccgac atgatgcaag acccggtcac ctccctgc acgcgtcaag 15120
 ttagcaacta cccgggtggtg ggcggccgac tcctgcccgt ctactccaag agcttcttca 15180
 acgagcaggc cgtctactcg cagcagctgc gcgccttcac ctgcgtcacg cacgtttca 15240
 accgcttccc cgagaaccag atccctgtcc gcccggccgc gcccaccatt accaccgtca 15300
 gtgaaaacgt tcctgcttc acagatcacg ggaccctgccc gctgcgcage agtatccggg 15360
 gagtccagcg cgtgaccgtc actgacgcca gacgcccac ctgccttac gtctacaagg 15420
 ccctggggcgt agtcgcgcgc cgcgtcttcc ctagccgcac ctctaaaaa atgtccattc 15480
 tcatctcgcc cagtaataac accgggttggg gcctgcgcgc gcccagcaag atgtacggag 15540
 ggcgcgcaca acgctccacg caacaccccg tgcgctgcg cgggcacttc cgcgtccct 15600
 ggggcgcctt caagggccgc gtgcgcgcgc gcaccaccgt ctagcgcgtg atcgaccagg 15660
 tggtggccga cgcgcgcac tacacgcccc cgcgcgcgc cgtctccacc gtggacgcgg 15720

tcatcgacag cgtggtgccc gacgcgcgcc ggtacgcccc caccaagagc cggcggcgcc 15780
gcatcgcccc gcgccacccgg agcacccccc ccatgcgcgc ggccgcagcc ttgctgcgca 15840
gggcaggcg cacgggacgc agggccatgc tcagggcgcc cagacgcgc gcctccggca 15900
gcagcagcgc cggcaggacc cgccagacgc cgccacggc ggccgcggcg gccatgcaca 15960
gcatgtcccc cccgcggcgc ggcaacgtgt actgggtgcg cgacgcgc accggtgtgc 16020
gcgtgccccgt ggcacccgc cccctcgca cttaagatg ctgacttcgc gatgtttagt 16080
tgtccccagcg gcgaggagga tgtccaagcg caaatacaag gaagagatgc tccaggtcat 16140
cgccctcgag atctacggcc cccgcggcgc ggtgaaggag gaaagaaaagc cccgcaaact 16200
gaagcgggtc aaaaaggaca aaaaggagga ggaagatgac ggactgggtgg agtttgtgc 16260
cgagttcgcc ccccgccggc gcgtgcagtgc gcgcggcgaa agtggaaac cggtgctgcg 16320
gccccggcacc acggtggtct tcacgccccgg cgagcggtcc ggctccgcct ccaagcgctc 16380
ctacgacgag gtgtacgggg acgaggacat cctcgagcag gcggtcgagc gtctggcg 16440
gtttgcgtac ggcaagcgca gcccggccgc gcccttggaa gaggaggcggttgtcc 16500
gctggaccac ggcaacccca cggcggccct gaagccgggtg accctgcagc aggtgttacc 16560
gagcgcggcg cccgcgggg gcttcaagcg cgagggcgcc gaggatctgt acccgaccat 16620
gcagctgatg gtggccaaagc gccagaagct ggaggacgtgc tgaggcaca tgaaggtgg 16680
ccccgaggtg cagcccggg tcaaggtgcg gccatcaag caggtggccc cggcctggg 16740
cgtgcagacc gtggacatca agatccccac ggagcccatg gaaacgcaga ccgagccgt 16800
gaagccccagc accagcacca tggaggtgca gacggatccc tggatgccag caccagttc 16860
caccagcaact cggcgaagac gcaagtacgg cggccgcgc ctgctgtgc ccaactacgc 16920
gctgcattcc tccatcatcc ccacgccccgg ctaccgcggc acgcgttctt accgcggcta 16980
caccagcage cggccggccgc agaccaccac cccgcggcggt cgtgcagcc gccgcagcag 17040
caccgcact tccgccttgg tgccggaggt gtatcgagc gggcgccgagc ctctgaccct 17100
gcccgcgcgc cgctaccacc cgacatcgcc catataacta cccgcctctta cttgcagata 17160
tggccctcac atgcgccttc cggcgtcccc ttacgggcta ccgaggaaga aagccgcgc 17220
gtagaaggct gacggggaaac gggctgcgtc gccatcacca cccgcggcg cgcgcacatca 17280
gcaagcggtt ggggggaggc ttccctggcccg cgctgttcc catcatgcgc gcccgcgc 17340
gggcgcatccc cggcatacg tccgtggcg tgcaggcctc tcagcgccac tgagacacaa 17400
aaaagcatgg atttgtataaaa aaaaaaaaaa tggactgacg ctccctggtcc tggatgtgt 17460

gttttagat ggaagacatc aattttcgta ccctggcacc gcgcacacggc acgcggccgt 17520
ttatggcac ctggagcgac atcggcaaca gccaactgaa cgggggcgcc ttcaattgga 17580
gcagtctctg gagcgggctt aagaatttcg ggtccacgct caaaaacctat ggcaacaagg 17640
cgtggAACAG cagcacaggg caggcgctga gggAAAAGCT gaaAGAACAG aacttccAGC 17700
agaagggttgt tgatggcctg gcctcaggca tcaacgggtt ggTTGACCTG gccaaccagg 17760
ccgtgcagaa acagatcaac agccgcctgg acgcggtccc gcccgcgggg tccgtggaga 17820
tgccccaggt ggaggaggag ctgcctcccc tggacaAGCG cggcgacaag cgaccgcgtc 17880
ccgacgcggA ggagacgctg ctgacgcaca cggacgagcc gccccgtac gaggaggcgg 17940
tgaAAACTGGG cctgcccacc acgcggcccg tggcgctct ggccacccgga gtgctgaaac 18000
ccagcagcag ccagccgcg accctggact tgccctcgcc tcgccccctcc acagtggcta 18060
agccccgtcc gccggtgcc gtcgcgtcgc gcgcgggggg aggccgcgggg caggcgaact 18120
ggcagagcac tctgaacAGC atcgtgggtc tggagtgca gagtgtaag cgccgcgcgt 18180
gctattaaaa gacactgttag cgcttaactt gcttgcgtgt gtgtatatgt atgtccgcgc 18240
accagaagga ggagtgtgaa gaggcgcgtc gcccgggtgc aagatggcca ccccatcgat 18300
gtgtccccag tggcggtaca tgcacatcgc cggacaggac gcttcggagt acctgagtcc 18360
gggtctggtg cagttcgccc gcgcacacaga cacctactc agtctgggg acaagtttag 18420
gaaccccacg gtggcgccca cgcacgatgt gaccacccgac cgcagccAGC ggctgacgct 18480
gcgccttcgtg cccgtggacc gcggggacaa cacctactcg tacaaagtgc gctacacgct 18540
ggccgtgggc gacaaccgcg tgctggacat ggcagcacc tactttgaca tccgcggcgt 18600
gctggaccgg ggccctagct tcaaacccta ctctggcacc gcctacaaca gcctagctcc 18660
caagggagct cccaaattcca gccagtgaaa gcaagcaaaa acaggcaatg ggggaactat 18720
ggaaacacac acatatggtg tggcccaat gggcgagag aatattacaa aagatggtct 18780
tcaaattgga actgacgtta cagcgaatca gaataaacca atttatgccc acaaaacatt 18840
tcaaccaAGAA ccgcaAGTAG gagaagaaaa ttggcaagaa actgaaaact tttatggcgg 18900
tagagctctt aaaaaagaca caaacatgaa accttgcgtat ggctcctatg cttagacccac 18960
caatgaaaaa ggaggtcaag ctaaactaa agttggagat gatggagttc caaccaaaga 19020
attcgacata gacctggctt tctttgatac tcccggtggc accgtgaacg gtcaagacga 19080
gtataaagca gacattgtca tgtataccga aaacacgtat ttggaaaactc cagacacgca 19140

tgtggtatac aaaccaggca aggatgatgc aagttctgaa attaacctgg ttcagcagtc 19200
tatgcccaac agacccaact acattgggtt cagggacaac tttatcggtc ttatgtacta 19260
caacagcaact ggcaatatgg gtgtgcttgc tggcaggcc tcccagctga atgctgtgg 19320
tgatttgc当地 gacagaaaca ccgagctgtc ctaccagctc ttgcttgaact ctttgggtga 19380
cagaaccggg tattttagta tgtggAACCA ggcgggtggac agttatgacc ccgatgtgcg 19440
catcatcgaa aaccatggtg tggaggatga attgccaaac tattgcttcc cttggacgg 19500
ctctggcaact aacgcccgt accaagggtgt gaaagaaaaa gatggtaaag atggtgatgt 19560
tgagagtgaa tggaaaaatg acgataactgt tgcaGCTCGA aatcaattat gtAAAGGTTAA 19620
catTTTCCGC ATGGAGATTA ATCTCCAGGC TAACCTGTGG AGAAGTTCC TCTACTCGAA 19680
cgtggccctg tacctgcccc actcctacaa gtacacgccc accaacgtca cgctggccgac 19740
caacaccaac acctacgatt acatgaatgg cagagtgaca cctccctcgc tggtagacgc 19800
ctaccccaac atcggggcgc gctggtcgct ggaccccatg gacaacgtca acccccttcaa 19860
ccaccaccgc aacgcgggccc tgcgttacccg ctccatgctc ctggcaacg ggctgtacgt 19920
GCCCTTCCAC ATCCAGGTGC CCCAAAAGTT TTTCGCCATC AAGAGCCTCC TGCTCCTGCC 19980
cgggtcc tac acctacgagt ggaacttccg caaggacgtc aacatgatcc tgcaGAGCTC 20040
cctaaggcaac gacctgcccga cggacggggc ctccatgccc ttccaccagca tcaacccctta 20100
cgccacccccc ttccccatgg cgcacaacac cgccctccacg ctgcaggcca tgctgccc 20160
cgacaccaac gaccaggcct tcaacgacta cctctggcg gccaacatgc tctaccccat 20220
cccggccaaac gccaccaacg tgcccatctc catccctcg cgcaactggg ccgccttccg 20280
cgatggtcc ttcacgccc tgaagacccg cgagacgccc tcgctggct ccgggttccg 20340
cccctacttc gtctactcg gctccatccc ctacctagac ggcacccctt acctcaacca 20400
caccccaag aaggcttcca tcacccctcgat cccctccgcg agctggcccg gcaacgaccg 20460
cctccctgacg cccaaacgagt tcgaaatcaa gcgcacccgtc gacggagagg gatacaacgt 20520
ggcccaagtgc aacatgacca aggactggtt cctggtccag atgctggccc actacaacat 20580
cggttaccag ggcttctacg tgcccgaggc ctacaaggac cgcatgtact ctttcccg 20640
caacttccag cccatgagcc gccagggtgt ggacgggtc aactacaagg actaccaggc 20700
cgtcaccctg gcctaccagc acaacaactc gggcttcgtc ggctacccctcg cgcccccac 20760
gcccaggccc cagccctacc ccgcaccaacta cccctacccg ctcatggca agagcgccgt 20820
cgccacgcgtc accccagaaaa agttccctcg cgacccgggtc atgtggcgcata ccccttctc 20880

cagcaacttc atgtccatgg gcgcgctcac cgacacctggc cagaacatgc tctacgcaa 20940
 ctccgcccac gcgctagaca tgaatttcga agtgcacccc atggatgagt ccacccttct 21000
 ctatgttgc ttgcagaatct tcgacgtcggt ccgagtgcac cagccccacc gggcgctcat 21060
 cgaagccgtc tacctgcgca cgccttctc ggccggcaac gccaccacct aagccgtct 21120
 tgcttcttgc aagatgacgg cgggctccgg cgagcaggag ctcagggcca tctccgcga 21180
 cctggctgc gggccctgct tcctggcac ctgcacaag cgctccctg gattcatggc 21240
 cccgcacaag ctggcctgcg ccatcgtaa cacggccggc cgcgagaccc gggcgagca 21300
 ctggctggcc ttgcgcctgaa acccgcgctc ccacacatgc tacctttcg acccccttcgg 21360
 gttctcgac gagcgcctca agcagatcta ccagttcgag tacgaggccc tgctgcgtcg 21420
 cagcgcctg gccaccgagg accgctgcgt caccctggaa aagtccaccc agaccgtgca 21480
 gggtccgcgc tggccgcctt gggggcttt ctgctgcatt ttcctgcacg cttcggtca 21540
 ctggccgcac cgccccatgg acaagaaccc caccatgaac ttactgacgg gggtgcccaa 21600
 cggcatgctc cagtcgcccc aggtggaaacc caccctgcgc cgcaaccagg aagcgtctta 21660
 ccgcatttcata aatgcggactt ccgcctactt tcgcgtccac cgcgccgcgc tcgagaaggc 21720
 caccgcatttc gaccgcatttgc atcaagacat gtaaaaaaacc ggtgtgtgtt tgtgaatgct 21780
 ttattcataa taaacacgcac atgtttatgc caccttcctt gaggctctga ctttatttag 21840
 aaatcgaagg gttctcgccg gctctcgca tggccgcgg gcagggatac gttgcggAAC 21900
 tggtaattgg gcagccactt gaactcgaaa atcagcagct tgggcacggg gaggtcgaaaa 21960
 aacgagtcgc tccacagctt gcgcgtgagt tgcagggcgcc caccgcggc gggcgccgg 22020
 atcttggaaat cgcagttggg acccgcggttc tgcgcgcgag agttgcggta cacggggttt 22080
 cagcacttggaa acaccatcag ggccgggtgc ttcaacgttg ccagcaccgt cgcgtcggt 22140
 atgccttcca cgtcccgatc ctcggcggttgc gccatcccgaa agggggatcat cttgcagg 22200
 tgccggccca tgctggcac gcagccgggc ttgtgggtgc aatcgacttgc cagggggatc 22260
 agcatcatct gggctgcgtc ggagctcatg cccgggtaca tggccatcat gaaagccctcc 22320
 agctggcgga aggccctgtcg cgccttgcgg ccctcggtga agaagacccc gcaggactt 22380
 ctagagaact ggttgggtggc gcagccggcg tcgtgcacgc agcagcgcgc gtcgttgg 22440
 ggcagctgca ccacgctgcg ccccccagcgg ttctgggtga tcttggcccg gttggggttc 22500
 tccttcagcg cgcgtgcggc gttctcgctc gccacatcca tctcgatagt gtgtcccttc 22560

tggatcatca cggccccgtg caggcacccgc agcttgcctt cggcttcgggt gcagccgtgc 22620
 agccacagcg cgccagccgggt gcactcccag ttcttgggg cgtatctggga gtgcgagtgc 22680
 acgaaggccct gcaggaagcg gcccatacgc gcggtcagggt tcttgggtgt ggtgaagggtc 22740
 agcgggatgc cgccgggtgc tcgcgttcaca tacaggtggc agatgcggcg gtacaccccg 22800
 ccctgctcggt gcatcagctg gaaggccggac ttccagggtcgc tctccacccgc gtaccggtec 22860
 atcagcagcg tcatcacttc catgcccttc tcccaaggccg aaacgatccgg caggctcagg 22920
 gggttttca ccgcattgtt catcttagtc gcccggccg aggtcagggg gtcgttctcg 22980
 tccagggtct caaacactcg ctgcgtcc ttctcgatga tgccgcacggg gggaaagctg 23040
 aagcccacgg ccgcagctc ctccctggcc tgccttcgtt cctcgctgtc ctggctgtatg 23100
 tcttgcggaa gcatatgtt ggttttgcgg ggttttttt tggggggcag aggccggccgc 23160
 gatgtgttgg gagagcgcga gttctcgatc accacgacta ttttttttc ttggccgtcg 23220
 tccgagacca cgcggccggta ggcattgcctc ttctggggca gaggccggagg cgacgggctc 23280
 tcgcgggtcg gccccgggtt ggcagagccc ctteccgcgtt cgggggtgcgtt ctcctggccgg 23340
 cgctgctctg actgacttcc tccgcggccg gccattgtgt tctccttaggg agcaacaaca 23400
 agcatggaga ctcaaggccatc gtcgcacaaca tcgcctatcg ccccccggccg caccggccac 23460
 gagaaccaggc agcagaatga aagcttaacc gccccggccg ccagccccac ctccgacgccc 23520
 gcggccccag acatgcaaga gatggaggaa tccatcgaga ttgacctggg ctacgtgacg 23580
 ccccgccggc acgaggaggaa gctggcagcg cgcttttcag ccccgaaaga gaaccaccaa 23640
 gagcagccag agcaggaagc agagaacgag cagaaccagg ctgggcacga gcatggccac 23700
 tacctgagcg gggcagagga cgtgctcatc aagcatctgg cccgcataatg catcatcgatc 23760
 aaggacgcgc tgctcgaccgc cggccgggtt cccctcagcg tggccggagct cagccgcgc 23820
 tacgagcgca accttttcgc gccgcgcgtt ccccccggc gccagccccaa cggcacctgt 23880
 gagcccaacc cgcgcctcaa ctttcttccgcgtt gtcgttgcgg tggccggagct cctggccacc 23940
 taccacccctt ttttcaagaa ccaaaggatc cccgtctctt gccgcgcctaa cccgcacccgc 24000
 gcccggccccc tgctcaaccc tggcccccggc gcccggccatc ctgatatcaca ctccctggaa 24060
 gaggttccca agatcttcga gggtctggc agcgcacgaga ctggggccgc gaacgtctg 24120
 caaggaagcg gagaggaggca tgagcaccac agcgcacccgc tggagttgga aggcgacaac 24180
 gcgccctgg cggccctcaa gccgcacggc gagctgaccc acttcgccta cccggccgtc 24240
 aacctgcccc ccaaggtcat gagcgcgcgtc atggaccagg tgctcatcaa gcgccctcg 24300

ccccttcgg aggaggagat gcaggacccc gagagttcgg acgagggcaa gcccgtggc 24360
 agcgacgagc agctggcgcg ctggctggga gcgagtagca ccccccagag cctgaaagag 24420
 cggcgcaagc tcatgatggc cgtggctctg gtgaccgtgg agctggagtg tctgcgcgc 24480
 ttctttcccg acgccccggac cctgcgcag gtcgaggaga acctgcacta cctttcagg 24540
 cacgggttcg tgcgccaggc ctgcaagatc tccaaacgtgg agctgaccaa cctggctcc 24600
 tacatggca tcctgcacga gaaccgcctg gggcaaaacg tgctgcacac caccctgcgc 24660
 ggggaggccc gccgcgacta catccgcgac tgcgtctacc tgcgtctcg ccacacctgg 24720
 cagacgggca tggcggtgtg gcagcagtgc ctggaggagc agaacctgaa agagctctgc 24780
 aagctcctgc agaagaacct caaggccctg tggaccgggt tgcacgagcg taccacccgc 24840
 tcggacctgg ccgacccat cttcccgag cgccgcggc tgacgtgcg caacgggctg 24900
 cccgacttta tgagccaaag catgttgc aaacttgcgt ctgcgtccct cgaacgcctc 24960
 gggatccctgc ccgcacccctg ctccgcgtcg ccctcgact tgcgtccct gacccctccgc 25020
 gagtgccccc cgccgcgtcg gagccactgc tacttgcgtgc gctggccaa ctacotggcc 25080
 taccactcgg acgtgatcga ggacgtcage ggcgagggtc tgctggagtg ccactgcgc 25140
 tgcaacccct gcacgcgcga ccgcctccctg gctgcaccc cccagctgtc gagcgagacc 25200
 cagatcatcg gcacccctcg gttgcaaggc cccggcgacg gcgaggccaa ggggggtctg 25260
 aaactcaccc cggggctgtg gacctcgcc tacttgcgcgca agttcgtgcc cgaggactac 25320
 catcccttcg agatcagggtt ctacgaggac caatcccagc cgcccaaggc cgagctgtcg 25380
 gcctgcgtca tcacccaggg ggccatccctg gccaattgc aagccatcca gaaatccgc 25440
 caagaatttc tgctgaaaaa gggccacggg gtctacttgg accccccagac cggagaggag 25500
 ctcaacccca gttcccccgg ggtgcggcc aggaagcagc aagaagctga aagtggagct 25560
 gccggccggc gaggatttgg aggaagactg ggagagcagt caggcagagg aggaggagat 25620
 ggaagactgg gacagcactc aggcagagga ggacagccctg caagacagtc tggaggagga 25680
 agacgagggtg gaggaggcag aggaagaagc agccgcggc agaccgtcg cctcgccgga 25740
 gaaagcaagc agcacggata ccacccctcg tccgggtcgg ggtcgccggc gccggccca 25800
 cagtaggtgg gacgagaccg ggcgcttccc gaacccacc acccagaccg gtaagaagga 25860
 gccggcaggga tacaagtcct ggccggggca caaaaacgc atcgctccct gcttgcaagc 25920
 ctgcggggc aacatctccct tcacccggcg ctacctgcgc ttccaccgcg gggtgaactt 25980

ccccccgaac atttgcatt actaccgtca cctccacagc ccctactact gtttccaaga 26040
 agaggcagaa acccagcagc agcagaaaaac cagcggcagc agcagctaga aaatccacag 26100
 cggcggcagg tggactgagg atcgcggcga acgagccggc gcagacccgg gagctgagga 26160
 accggatctt tcccacccctc tatgccatct tccagcagag tcggggcag gagcaggaac 26220
 taaaagtcaa gaaccgttct ctgcgctcgc tcaccccgag ttgtctgtat cacaagagcg 26280
 aagaccaact tcagcgcact ctcgaggacg ccgaggctct cttcaacaag tactgcgcgc 26340
 tcactcttaa agagtagccc ggcggccccc acacacggaa aaaggcggga attacgtcac 26400
 cacctgcgcc cttcgccccga ccatcatgag caaagagatt cccacgcctt acatgtggag 26460
 ctaccagccc cagatgggcc tggccggcccg cgccgcccag gactactcca cccgcatgaa 26520
 ctggctcagt gcccggcccg cgatgatctc acgggtgaat gacatccgcg cccaccgaaa 26580
 ccagatactc ctagaacagt cagcgatcac cgccacgccc cgccatcacc ttaatccgcg 26640
 taattggccc gccgccccctgg tgtaccagga aatccccag cccacgaccg tactacttcc 26700
 gcgagacgcc caggccgaag tccagctgac taactcaggt gtccagctgg cccggccgc 26760
 cgccctgtgt cgtcaccggc cccgtcaggg tataaagcgg ctggtgatcc gaggcagagg 26820
 cacacagctc aacgacgagg tggtgagctc ttcgctgggt ctgcgacctg acggagtctt 26880
 ccaactcgcc ggatcgggga gatcttcctt cccgtcgtt cagggcgtcc tgactttgg 26940
 gagtttgtcc tcgcagcccc gtcggggcgg catggcaact ctccagttcg tggaggagtt 27000
 cactccctcg gtctacttca accccttctc cggctcccc ggcactacc cggacgagtt 27060
 catcccgAAC ttcgacgcca tcagcgagtc ggtggacggc tacgattgaa tgtcccatgg 27120
 tggcgcagct gacctagctc ggcttcgaca cctggaccac tgccgcccgt tccgctgttt 27180
 cgctcgggat ctcgcccagt ttgcctactt tgagctgccc gaggagcacc ctcagggccc 27240
 agccccacgga gtgcggatca tcgtcgaagg gggcctcgac tcccacctgc ttccggatctt 27300
 cagccagcga ccgatcctgg tcgagcgcga acaaggacag acccttctta ctttgtactg 27360
 catctgcaac caccggccc tgcattaaag tctttgtgt ctgctgtgtatctgactt 27420
 taaaagctga gatcagcgcac tactccggac tcgattgtgg tggccctgtatctgactt 27480
 ccctgttctt caccgggaac gagaccgagc tccagctcca gtgtaaagccc cacaagaagt 27540
 acctcacctg gtcgttccag ggctccccga tcgcccgtgtt caaccactgc gacaacgacg 27600
 gagtcctgtatctgactt gacccggccct gccaacctta cttttccac ccgcagaagc aagtcacac 27660
 tttccaaacc ctccctcccc gggaccttac agtgcgtctc aggaccctgc catcacac 27720

tccacactgat	cccgaataacc	acagcgccgc	tcccccgtac	taacaaccaa	actacccacc	27780
aacggcaccg	tcgcgacctt	tcctctgaat	ctaataccac	taccggaggt	gagctccgag	27840
gtcgaccaac	ctctgggatt	tactacggcc	cctgggaggt	ggtgggggta	atagcgctag	27900
gcctagttgc	gggtgggctt	ttgggtctct	getacctata	cctcccttgc	tgttcgtact	27960
tagtggtgct	gtgttgcgtgg	tttaagaata	gggaaagatc	accctagtga	gctgcgggtc	28020
gctggtggcg	gtgttgcgttt	cgattgtggg	actgggcccgc	gcggctgttag	tgaaggagaa	28080
ggccgatccc	tgcttgcatt	tcaatccaa	caaatgccag	ctgagtttgc	agcccgatgg	28140
caatcggtgc	gcggtaactga	tcaagtgcgg	atggaaatgc	gagaacgtga	aatcgagta	28200
caataacaag	actcggaaaca	atactctcgc	gtccgtgtgg	cagccccggg	accccgagtgc	28260
gtacaccgtc	tctgtccccg	gtgctgacgg	ctccccgcgc	accgtgaata	atactttcat	28320
ttttgcgcac	atgtgcaaca	cggtcgttg	gatgagcaag	cagtacgata	tgtggccccc	28380
cacgaaggag	aacatcgtgg	tcttctccat	cgcttacagc	ctgtgcacgg	cgctaattcac	28440
cgctatcggt	tgcctgagca	ttcacatgt	catcgctatt	cgcggcagaa	ataatgcgca	28500
gaaagagaaaa	cagccataaac	acgttttttc	acacaccttg	tttttacaga	aatgcgtct	28560
gttaaatttt	ttaaacattt	tgctcagtat	tgcttatgcc	tctggttatg	caaacataca	28620
gaaaaccctt	tatgttaggat	ctgatggtag	actagagggt	acccaatcac	aagccaaagg	28680
tgcatggtat	ttttatagaa	ccaacactga	tccagttaaa	ctttgttaagg	gtgaattgcc	28740
gcgtacacat	aaaactccac	ttacatttag	ttgcagcaat	aataatctta	cacttttttc	28800
aattacaaaa	caatataactg	gtacttatta	cagtacaaac	tttcatacag	gacaagataa	28860
atattataact	gttaaggttag	aaaatccat	cactcctaga	actaccacca	ccaccactac	28920
tgcaaaagccc	actgtgaaaa	ctacaactag	gaccaccaca	actacagaaa	ccaccaccag	28980
cacaacactt	gctgcaacta	cacacacaca	caactaagcta	acttacaga	ccactaatga	29040
tttgatcgcc	ctgctgcaaa	agggggataa	cagcaccact	tccaatgagg	agatacccaa	29100
atccatgatt	ggcattattt	ttgctgtat	ggtgtgcatg	ttgatcatcg	ccttgcgtat	29160
ggtgtactat	gccttctgtat	acagaaagca	cagactgaac	gacaagctgg	aacacttact	29220
aagtgttcaa	tttaattttt	ttagaaccat	gaagatccta	ggccttttta	gttttctat	29280
cattacctct	gctctttgtg	aatcagtgga	tagagatgtt	actattacca	ctggttctaa	29340
ttatacactg	aaaggccac	cctcaggtat	gcttcgtgg	tattgctatt	ttggaactga	29400

cactgatcaa actgaattat gcaatttca aaaaggcaaa acctcaaact ctaaaatctc 29460
taattatcaa tgcaatggca ctgatctgat actactaat gtcacgaaaag catatggtgg 29520
cagttattat tgccctggac aaaacactga agaaatgatt tttacaaaag tggaagtgg 29580
tgatcccact acaccaccca ccaccacaac tattcatacc acacacacag aacaaacacc 29640
agaggcaaca gaagcagagt tggccttcca ggtcacgga gattccttgc ctgtcaatac 29700
ccctacaccc gatcagcggt gtccggggcc gctagtcagc ggcattgtcg gtgtgcttc 29760
gggatttagca gtcataatca tctgcatgtt cattttgtt tgctgtata gaaggctta 29820
ccgacaaaaa tcagacccac tgctgaacct ctatgtttaa tttttccag agccatgaag 29880
gcagttagcg ctctagttt ttgttcttg attggcattt gttttatag taaaattacc 29940
agagttagct ttatcaaaca tgttaatgtt actgaaggag ataacatcac actagcaggt 30000
gtagaaggtg ctcaaaacac cacctggaca aaataccatc taggatggag agatattgc 30060
acctggaatg taacttatta ttgcataatggaa gttaatctta ccattgttaa cgctaaccaa 30120
tctcagaatg gtttaattaa aggacagagt gtttgtgtga ccagtgtgg gtactatacc 30180
cagcatagtt ttaactacaa cattactgtc ataccactgc ctacgcctag cccacctagc 30240
actaccacac agacaaccac atacagtaca tcaaattcgc ctaccacccac tacagcagca 30300
gaggttgcgc gctcgctgg ggtccgagtg gcattttga tggggcccc atctagcagt 30360
cccactgcta gtaccaatga gcagactact gaattttgt ccactgtcgag gagccacacc 30420
acagctacccctt ccagtgcctt ctctagcacc gccaatctct cctcgcttc ctctacacca 30480
atcagccccg ctactactcc tagccccgtt ccttccca cttccctgaa gcaaacagac 30540
ggcggcatgc aatggcagat caccctgctc attgtgatcg ggttggtcat cctggccgtg 30600
ttgctctact acatcttctg ccggccgatt cccaaacgcgc accgcaagcc ggcctacaag 30660
cccatcgta tcgggcagcc ggagccgctt caggtggaaag ggggtctaag gaatcttetc 30720
ttctttta cagttatggtg attgaactat gattcctaga caattttga tcactattct 30780
tatctgcctc ctccaaagtct gtgccaccctt cgctctggtg gccaacgcgc gtccagactg 30840
tattggggccc ttgccttcctt acgtgcctt tgccctcgatc acctgcattt gctgtgttag 30900
catagtctgc ctgcttatca ctttcttcca gttcattgac tggatcttgc tgccatcg 30960
ctacactgc caccacccccc agtaccgcga ccagcgagtg ggcgcagctgc tcaggctcc 31020
ctgataagca tgccggctct gctacttctc ggccttctgc tggtagtgc ccccccgtccc 31080
gtcgacccccc ggtccccccac tcagtcccccc gaggaggttc gcaaattgcaaa attccaagaa 31140

ccctggaaat tcctcaaatg ctaccgccaa aaatcagaca tgcatacccag ctggatcatg 31200
atcattggga tcgtgaacat tctggcctgc accctcatct cctttgtat ttacccctgc 31260
tttgactttg gttggaaatc gccagaggcg ctctatctcc cgccctgaacc tgacacacca 31320
ccacagcagc aacctcaggc acacgcacta ccaccaccac agccttaggcc acaatacatg 31380
ccccatattag actatgaggc cgagccacag cgacccatgc tccccgtat tagttacttc 31440
aatctaaccg gcggagatga ctgaccact ggccaataac aacgtcaacg accttctcct 31500
ggacatggac ggccgcgcct cggagcagcg actcgcccaa cttcgcattc gtcagcagca 31560
ggagagagcc gtcaaggagc tgcaggacgg catagccatc caccagtgca agagaggcat 31620
cttctgcctg gtgaaacagg ccaagatctc ctacgaggtc acccagacgg accatgcct 31680
ctcctacgag ctcctgcagc agcgcacagaa gttcacctgc ctggtcggag tcaacccat 31740
cgtcatcacc cagcagtcgg gcgataccaa ggggtgcate cactgctct gegactcccc 31800
cgactgcgtc cacactctga tcaagacccct ctgcggcctc cgccgacctcc tccccatgaa 31860
ctaattcaccc ccttatccag taaaataaaat atcatattga tgatgattta aataaaaaaa 31920
ataatcattt gatttggaaat aaagatacaa tcataattgat gatttggatgtaa aaaaaaaaaat 31980
aaagaatcac ttacttgaaa tctgataacca ggtctctgtc catgtttctt gccaacacca 32040
cctcactccc ctcttcccg ctctggtaact gcaaggccccg gcgggctgca aacttccctcc 32100
acacgctgaa ggggatgtca aattccctct gtcctcaat cttcattttt tcttctatca 32160
gatgtccaaa aagcgcgtcc gggatggatga tgacttcgac cccgtctacc cctacgatgc 32220
agacaacgca ccgaccgtgc ctttcatcaa ccccccttc gtctttagt atggattcca 32280
agagaagccc ctgggggtgt tgccctgtcg actggctgac cccgtcacca ccaagaacgg 32340
ggaaatcacc ctcaagctgg gagaggggggt ggacctcgac tcgtcgggaa aactcatctc 32400
caacacggcc accaaggccc ccgcctct cagtttca aacaacacca tttcccttaa 32460
aactgctgcc ctttctaca acaacaatgg aactttaagc ctcaatgtct ccacaccatt 32520
agcagtattt cccacattta aactttagg cataagtctt ggaaacggc ttcagacttc 32580
aaataagttt ttgactgtac aactaactca tccttctaca ttcaatgtca atagcatcac 32640
agtaaaaaaca gacaaaggac tatatatattaa ctccagtgaa aacagaggac ttgaggctaa 32700
tataaggccta aaaagaggac tagtttttga cggtaatgtt attgcaacat atattggaaa 32760
tggcttagac tatggatctt atgatagtga tggaaaaaca agacccgtaa ttacccaaat 32820

tggagcagga ttaaattttg atgctaacaa agcaatagct gtcaaactag gcacaggaaa 32880
aagttttgc tccgctggtg cttgacagc tgaaaacaaa caggatgaca agctaacact 32940
ttggactacc cctgacccaa gccctaattt tcaattactt tcagacagag atgccaattt 33000
. tactctctgt cttacaaaat gcggtagtca aatacttaggc actgtggcag tggcggctgt 33060
tactgttagga tcaagcactaa atccaattaa tgacacagtc aaaagcgcac tagtttcc 33120
tagatttgat tccgatggtg tactcatgtc aaactcatca atggtaggtg attactggaa 33180
cttttagggag ggacagacca ctcaaagtgt agcctataca aatgctgtgg gattcatgcc 33240
aaatataggc gcatatccaa aaacccaaag taaaacaccc taaaatagca tagtcagtca 33300
ggtatattta actggagaaaa ctactatgcc aatgacacta accataactt tcaatggcac 33360
tcatgaaaaa gacacaaccc cagttagcac ctactctatg acttttacat ggcagtggac 33420
tggagactat aaggacaaaa atattaccc ttgttaccaac tcattctt tttccatcat 33480
cgcccaggaa taatcccacc cagcaagcca accccttttc ccaccaccc ttgtctatatg 33540
gaaactctga aacagaaaaa taaagttcaa gtgtttatt gaatcaacag ttttacagga 33600
ctcgagcagt tattttccct ccaccctccc aggacatgga atacaccacc ctctcccccc 33660
gcacagcctt gaacatctga atgccattgg tcatggacat gcttttggtc tccacgttcc 33720
acacagtttc agagcgagcc agtctcgat cggcaggaa gatgaaaccc tccggcact 33780
cccgcatctg cacctcacag ctcaacagct gaggattgtc ctccgggttc gggatcacgg 33840
ttatctggaa gaagcagaag agccgggtg ggaatcatag tccgcaacg ggatcggccg 33900
gtgggtgcgc atcaggcccc gcagcgtcg ctggccggc cgctccgtca agctgctgt 33960
cagggggttc ggttccaggg actccctcag catgatgcc acggccctca gcatcagtcc 34020
tctggtgcgg cgggcgcagc agcgcattcg aatctcgctc aggtcaactgc agtacgtgca 34080
acacaggacc accaggttgt tcaacagtc atagttcaac acgctccagc cgaaactcat 34140
cgccggaaagg atgctacccca cgtggccgtc gtaccagatc ctccaggtaaa tcaagtggcg 34200
ctccctccag aagacgctgc ccatgtacat gatctcttgc ggcattgtggc ggttcaccac 34260
ctcccggtac cacatcaccc tctggttgaa catgcagccc cggatgatcc tgcggaaacca 34320
cagggccagc accgccccgc cggccatgca gcaagagac cccggatccc ggcaatgaca 34380
atggaggacc caccgctcgat acccggtggat catctgggag ctgaacaagt ctatgttggc 34440
acagcacagg catatgctca tgcattctt cagcactctc agctccctcg gggtcaaaac 34500
catatcccag ggcacggggaa actcttgcag gacagcgaac cccgcagaac agggcaatcc 34560

tcgcacataa cttacattgt gcatggacag ggtatcgcaa tcagggcagca ccgggtgatc 34620
ctccaccaga gaagcgccgg tctcggtctc ctcacagcgt ggttaaggggg ccggccgata 34680
cgggtgatgg cgggacgcgg ctgatcggt tctcgaccgt gtcatgatgc agttgttgc 34740
ggacattttc gtacttgctg tagcagaacc tggtcgggc gctgcacacc gatcgccggc 34800
ggcgggtctcg ggcgttggaa cgctcggtgt taaagttgtaa aacagccac tctctcagac 34860
cgtgcagcag atcttagggcc tcaggagtga tgaagatccc atcatgcctg atagctctga 34920
tcacatcgac cacccgtggaa tggccaggc ccagccagat gatgcaattt tggtgggtt 34980
cggtgacggc gggggagggaa agaacaggaa gaaccatgtat taactttaa tccaaacggt 35040
ctcggagcac ttcaaaatga aggtcacggaa gatggcacct ctcgcccccg ctgtgttgg 35100
ggaaaataac agccaggtca aaggtgatac gggtctcgag atgttccacg gtggcttcca 35160
gcaaagcctc cacgcgcaca tccagaaaca agacaatagc gaaagcggga gggttctcta 35220
attcccaac catcatgtta cactcctgca ccattccccag ataattttca tttttccagc 35280
cttgaatgtatc tcgaactagt tcctgaggtaa aatccaagcc agccatgata aaaagctcgc 35340
gcagagcacc ctcccacccggc attcttaagc acaccctcat aattccaaga tattctgctc 35400
ctgggttccacc tgacagcagat tgacaagcgg aatatcaaaa tctctgccgc gatccctgag 35460
ctcctccctc agcaataact gtaagtactc tttcatatcg tctccgaaat ttttagccat 35520
aggaccccca ggaataagag aaggcaagc cacattacag ataaaccgaa gtccccccca 35580
gtgaggcattt ccaaattgtaa gattgaaata agcatgtgg cttagccgg tgatatcttc 35640
cagataactg gacagaaaaat cgggtaaagca atttttaaga aaatcaacaa aagaaaaatc 35700
ttccaggtgc acgttttaggg cctcgggaac aacgatggag taagtgcag gggtgcttc 35760
cagcatggtt agttagctga tctgtaaaaaa aacaaaaaat aaaacattaa accatgctag 35820
cctggcgaac aggtgggtaa atcggttctct ccagcaccag gcaggccacg gggctccgg 35880
cgccgaccctc gtaaaaatttgc tgcgtatgtat tgaaaaccat cacagagaga cggtccgg 35940
ggccggcgtg aatgattcga gaagaagcat acaccccccgg aacattggag tccgtgagtg 36000
aaaaaaaaacgc gccgaggaag caatgaggca ctacaacgct cactctcaag tccagcaaag 36060
cgatgccatg cgatgtaaacgca aaaaaattttt caggtgcgtat aaaaatgtaa ttactccct 36120
cctgcacagg cagcgaagct cccgatccct ccagatacac atacaaagcc tcagcgtcca 36180
tagcttaccg agcggcagca gcagcggcac acaacaggcg caagagtcaag agaaaaagact 36240

gagctctaac ctgtccgccc gtcctctgct caatatatacg ccccagatct acactgacgt 36300
 aaaggccaaa gtctaaaaat acccgccaaa taatcacaca cggccagcac acgcccagaa 36360
 accgggtgaca cactcagaaa aatacgcgca cttcctcaaa cggccaaact gccgtcattt 36420
 ccgggttccc acgctacgta atcaaaacac gactttcaaa ttccgtcgac cgtaaaaac 36480
 atcacccgac ccgccccctaa cggtcggcgc tccccgagcc aatcaccttc ctccctcccc 36540
 aaattcaaac agtcatttg catattaacg cgcaccaaaa gttttaggta tattattgat 36600
 gatg 36604

<210> 3
 <211> 36535
 <212> DNA
 <213> chimpanzee adenovirus serotype Pan7

<400> 3
 catcatcaat aatatacctc aaacttttgg tgccgttaa tatgcaaatg agctgttga 60
 atttggggag ggaggaaggt gattggccga gagacggcg accgttaggg gcggggcggg 120
 tgacgtttt aatacgtggc cgtgaggcg agccggtttgc aagttctcg tgggaaaagt 180
 gacgtcaaac gaggtgtggt ttgaacacgg aaatactcaa tttcccgcg ctctctgaca 240
 gaaaaatgagg tgtttctggg cggatgcaag tgaaaacggg ccattttcgc gcgaaaactg 300
 aatgaggaag tgaaaatctg agtaatttcg cggttatggc agggaggagt atttgccgag 360
 ggccgagtag actttgaccg attacgtggg gtttcgatt accgtattt tcacctaatt 420
 ttccgcgtac ggtgtcaaaag tccgggtttt ttacgttaggc gtcagctgat cgccagggt 480
 tttaaacctg cgctctctag tcaagaggcc actcttgagt gccagcgagt agagtttct 540
 cctccgcgcgc gcgagtcaga tctacactt gaaagatgag gcacctgaga gacctgccc 600
 gtaatgtttt cctggctact gggAACGAGA ttctggattt ggtgggtggac gccatgatgg 660
 gtggcgaccc tcctgagcccc cctacccat ttgaggcgcc ttcgtgtac gatggat 720
 atctggaggt ggatgtgccc gagaacgacc ccaacgagga ggcgggtgaat gatggat 780
 gcgatgcgcgc gctgtggct gccgagcagg ctaatacgga ctctggctca gacagcgatt 840
 cctctctcca taccccgaga cccggcagag gtgagaaaaa gatccccgag cttaaagggg 900
 aagagctcga cctgcgtgc tatgaggaat gcttcctcc gagcgatgat gaggaggacg 960
 aggaggcgat tcgagctgca tcgaaccagg gagtgaaagc tgcgggcgaa agcttagcc 1020
 tggactgtcc tactctgccc ggacacggct gtaagtcttgc tgaatttcat cgcatgaata 1080

ctggagataa	gaatgtgatg	tgtgccctgt	gctatatatgag	agcttacaac	cattgtgttt	1140
acagtaagtg	tgattaactt	tagttggaa	ggcagagggt	gactgggtgc	tgactgggttt	1200
atttatgtat	atgtttttt	atgttaggt	cccgctctcg	acgttagatga	gacccccact	1260
tcagagtgc	tttcatcacc	cccagaaatt	ggcgaggaac	cgccccaga	tattattcat	1320
agaccagttg	cagtgagagt	caccggggcg	agagcagctg	tggagagttt	ggatgacttg	1380
ctacagggtg	gggatgaacc	tttggacttg	tgtacccgga	aacgccccag	gcactaagtg	1440
ccacacatgt	gtgtttactt	aaggtgatgt	cagtatttat	agggtgttga	gtgcaataaa	1500
atccgtgtt	actttaagtg	cgtggtttat	gactcagggg	tggggactgt	gggtatataa	1560
gcaggtgcag	acctgtgtgg	tcagttcaga	gcaggactca	tggagatctg	gacggctttg	1620
gaagacttcc	accagactag	acagctgcta	gagaactcat	cgagggggt	ctcttacctg	1680
tggagattct	gttcgggtgg	gcctctagct	aagctagtct	atagggccaa	acaggattat	1740
aaggatcaat	ttgaggatat	tttgagagag	tgtcctggta	ttttgactc	tctcaacttg	1800
ggccatcagt	ctcactttaa	ccagagtatt	ctgagagccc	ttgactttc	tactcctggc	1860
agaactaccg	ccgcggtagc	ctttttgcc	tttacccctg	acaaatggag	tcaagaaacc	1920
catttcagca	gggattaccg	tctggactgc	ttagcagtag	ctttgtggag	aacatggagg	1980
tgccagcgcc	tgaatgcaat	ctccggctac	ttgccagtagc	agccggtaga	cacgctgagg	2040
atcctgagtc	tccagtcacc	ccaggaacac	caacgcgc	agcagccgca	gcaggagcag	2100
cagcaagagg	aggaggagga	tgcagaagag	aacccgagag	ccggctctgga	ccctccgggt	2160
gcggaggagg	aggagtagct	gacttggttc	ccgagctgcg	ccgggtgctg	actaggtctt	2220
ccagtgacg	ggagaggggg	attaagcggg	agaggcatga	ggagactagc	cacagaactg	2280
aactgactgt	cagtctgatg	agccgcaggc	gcccagaatc	ggtgtgggtgg	catgaggttc	2340
agtcgcaggg	gatagatgag	gtctcggtga	tgcatacgaa	atattccctg	gaacaagtca	2400
agactgttg	gttggaggct	gaggatgatt	gggaggtagc	catcaggaat	tatgccaagc	2460
tggctctgaa	gccagacaag	aagtacaaga	ttaccaaact	gattaatatc	agaaattcct	2520
gctacatttc	aggaaatggg	gccgagggtgg	agatcagtagc	ccaggagagg	gtggccttca	2580
gatgttgtat	gatgaatatg	tacccgggggg	tggtgggcat	ggagggagtc	acctttatga	2640
acgcgagggt	caggggtgat	gggtataatg	gggtgggttt	tatggccaac	accaagctga	2700
cagtgcacgg	atgctcccttc	tttgggttca	ataacatgtg	catcgaggcc	tggggcagtg	2760
tttcagtgag	gggatgcagc	ttttcagcca	actggatggg	ggtcgtgggc	agaaccaaga	2820

gcaagggtgtc	agtgaagaaa	tgcctgttcg	agaggtgccca	cctgggggtg	atgagcgagg	2880
gcgaagccaa	agtcaaacac	tgcgcctcta	ctgagacggg	ctgctttgtg	ctgatcaagg	2940
gcaatccccaa	agtcaagcat	aacatgatct	gtggggcctc	ggatgagcgc	ggctaccaga	3000
tgctgacctg	cgccggtggg	aacagccata	tgctgccac	cgtcatgtg	acctcgacc	3060
cccgcaagac	atggcccag	ttcgagcaca	acgtcatgac	ccgatgcaat	gtgcacctgg	3120
ggtcccggcg	aggcatgttc	atgcctacc	agtcaacat	gcaatttgtg	aaggtgctgc	3180
tggagccccg	tgccatgtcc	agagtgagcc	tgacgggggt	gtttgacatg	aatgtggagc	3240
tgtggaaaat	tctgagatat	gatgaatcca	agaccaggta	ccgggcctgc	gaatgcggag	3300
gcaagcacgc	caggcttcag	cccgtgtgt	tggaggtgac	ggaggacctg	cgaccccgatc	3360
atttggtgtt	gtccctgcaac	gggacggagt	tcggctccag	cgggaaagaa	tctgactaga	3420
gtgagtagtg	tttgggggag	gtggagggt	tgtatgaggg	gcagaatgac	taaaatctgt	3480
gtttttctgt	gtgttgcagc	agcatgagcg	gaagcgccctc	ctttgaggga	ggggtattca	3540
gccttatct	gacggggcgt	ctccctctt	gggcgggagt	gcgtcagaat	gtgatggat	3600
ccacgggtgga	cggccggccc	gtcagcccg	cgaactcttc	aaccctgacc	tacgcgaccc	3660
ttagctccctc	gtccgtggac	gcagctgccg	ccgcagctgc	tgcttccgc	gcgcagcggc	3720
tgcgccgaat	ggccctggc	gccggctact	acagctctct	ggtggccaac	tgcacttcca	3780
ccaataatec	cgccagectg	aacgaggaga	agctgctgt	gctgatggcc	cagctcgagg	3840
ccctgaccca	gcgcctggc	gagctgaccc	agcaggtggc	tcagctgcag	gcggagacgc	3900
gggcccgcgt	tgccacggtg	aaaaccaaat	aaaaaatgaa	tcaataaaata	aacggagacg	3960
gttggattt	ttaacacaga	gtcttgaatc	tttatttgc	ttttcgccgc	cggttaggccc	4020
tggaccacccg	gtctcgatca	ttgagcaccc	ggtggatttt	ttccaggacc	cggttagaggt	4080
gggcttggat	gttggaggtac	atgggcatga	gcccgtcccc	gggggtggagg	tagtccatt	4140
gcagggccctc	gtgtcgaaaa	gtgggtttgt	aaatcaccca	gtcatagcag	gggcgcaggg	4200
cgtggtgcgt	cacgatgtcc	ttgaggagga	gactgatggc	cacggcagc	cccttgggt	4260
agggtttgac	gaacctgttg	agctggagg	gatgcatgcg	gggggagatg	agatgcacat	4320
tggccctggat	ctttagattt	gcatgttcc	ccccagatc	ccggccggggg	ttcatgttgc	4380
gcaggaccac	cagcacggtg	tatccggtgc	acttggggaa	tttgcacatgc	aacttggaaag	4440
ggaaggcgtg	aaagaatttgc	gagacgcct	tgtgacccgc	caggtttcc	atgcactcat	4500

ccatgatgat	ggcgatggc	ccgtggcgg	cggcctggc	aaagacgtt	cgggggtcg	4560
acacatcgta	gttgtggtcc	tgggtgagct	cgtcataggc	catttaatg	aatttgggc	4620
ggagggtgcc	cgactgggg	acgaagggtgc	cctcgatccc	gggggcgtag	ttccccctcgc	4680
agatctgcat	ctcccaggcc	ttgagctcg	agggggggat	catgtccacc	tgcggggcga	4740
tgaaaaaaaaac	ggtttccggg	gcgggggaga	tgagctggc	cgaaagcagg	ttccggagca	4800
gctgggactt	gccgcagccg	gtggggcgt	agatgacccc	gatgaccggc	tgcaagggtggt	4860
agttagggga	gagacagctg	ccgtcctcgc	ggaggagggg	ggccacctcg	ttcatcatct	4920
cgcgcacatg	catgttctcg	cgcacgagtt	ccgcccaggag	gcgcctcgccc	cccagcgaga	4980
ggagcttttg	cagcgaggcg	aagttttca	gcggctttag	yccgtcgcc	atgggcattt	5040
tggagagggt	ctgttgcaag	agttccagac	ggtcccagag	ctcggtgatg	tgctctaggg	5100
catctcgatc	cagcagacct	cctcgtttcg	cggttgggg	cgactgcggg	atagggcac	5160
cagggcgtgg	gcgtccagcg	aggccagggt	ccggtccttc	cagggtcgca	gggtccgcgt	5220
cagcgtggc	tccgtcacgg	tgaaggggtg	cgccggggc	tgggcgttgc	cgagggtgcg	5280
cttcaggctc	atccggctgg	tcgagaaccg	ctcccggtcg	gcgcctcg	cgtcgccag	5340
gtagcaattt	agcatgagtt	cgtagtttag	cgccctcgcc	gcgtggccct	tggcgccgag	5400
cttacctttt	gaagtgtgtc	cgcagacggg	acagaggagg	gactttaggg	cgtagagctt	5460
gggggcgagg	aagacggact	cgggggcgta	ggcgtcccg	ccgcagctgg	cgcagacggt	5520
ctcgcaactcc	acgagccagg	tgaggtcg	ccgggttgggg	tcaaaaacga	gttttctcc	5580
gtgtttttt	atgcgtttct	tacctctgg	ctccatgago	tcgtgtcccc	gctgggtgac	5640
aaagaggctg	tccgtgtccc	cgtagaccga	ctttatggc	cggcctcga	gcggggtgcc	5700
gcggcctcg	tcgttagagga	accccgccca	ctccgagacg	aaggcccggg	tccaggccag	5760
cacgaaggag	gccacgtggg	agggtagcg	gtcggtgtcc	accagcggt	ccaccttctc	5820
caggtatgc	aagcacatgt	ccccctcg	cacatccagg	aaggtgattt	gctttaagt	5880
gtaggccacg	tgaccggggg	tccggccgg	gggggtataa	aaggggcg	gcgcctgc	5940
gtcctcaactg	tctccggat	cgtgtccag	gagcgc	tgttgggt	ggtattccct	6000
ctcgaaggct	ggcataaacct	cgcaactcg	gtgtcagtt	tctagaaacg	aggaggattt	6060
gatattgacg	gtgcgttgg	agacgcctt	catgagcccc	tcgtccatct	ggtcagaaaa	6120
gacgatctt	ttgttgcga	gttgggtggc	gaaggagccg	tagaggcgt	tggagaggag	6180
cttggcgtatg	gagcgcatgg	tctggttttt	ttccctgtcg	gcgcgtcct	tggcggcgat	6240

gtttagctgc acgtactcgc	6300
gcccacgc cttccatcg	6360
ggaaagacgg tggtagctc	6420
gtcggcacg attctgaccc	6480
gccagccgcg gttgtcagg	6540
gtatgaggt ccacgcttgt	6600
ggccacctcg cccgcgcagg	6660
gctcggttgt ccacgagg	6720
ccccccct tgcgcgagca	6780
gaaggggggc agcgggtcca	6840
gcatgagctc gtccgggggg	6900
tccggcgttcca cggtgaagat	6960
gcccggcaga agctcggtgt	7020
cgaagtagct gatgcagg	7080
tccagatcg tccagatcg	7140
ccagcgccgc	7200
ttggcgttgc ggggttc	7260
gggtgtcg agcgtcgagg	7320
cgtacatgcc gcagatgtcg	7380
tagacgtaga ggggtcc	7440
gaggacgccc atgttaggtgg	7500
ggtagcagcg cccccccgg	7560
atgctggcgc gcactagtc	7620
gtacagctcg tgcgagggcg	7680
cgaggagccc cgtgcgcagg	7740
ttggagcgtt ggggttttc	7800
ggcgcgttag acgatctggc	7860
ggaagatggc gtggagttg	7920
gaggagatgg tgggcctctg	
gaagatgttg aagtggcgt	
ggggcaggcc gaccgagtcc	
ctgatgaagt gggcgttagga	
gtcctgcagc ttggcgtacga	
gctcggtgt gacgaggacg	
tccagggcgc agtagtcgag	
ggtcttttgg atgatgtcg	
acttgagctg gcccctctgc	
tccacacgt cgcggttgag	
aaggaactct tcgcggtcct	
tccagtactc ttcgaggggg	
aaccctgtct gatcgac	
gtaagagccc accatgtaga	
actggttgac ggccttgtag	
gcccgttccac gggggggcg	
taagcttgcg cggccttgcg	
caggaggtg tgggtgaggg	
cgaaggtgtc	
gcccaccatg accttgcg	
actggtgctt gaagtcgagg	
tcgtcgacgc cgccctgtc	
ccagagctgg aagtccgtgc	
gttctttgtt ggcgggttg	
ggcaaagcga aagtaacatc	
gttgaagagg atcttgcgg	
cgcggggcat gaagttgcga	
gtgatgcgga aaggctgggg	
cacctcgccc cggttgttga	
tgacctggc ggcgaggacg	
atctcgatcg agetcgttat	
gttgcgtcccg acgatgtaga	
gttccacgaa tcgcggcgg	
cccttaacgt gggcagctt	
cttgcgtcg tcttaggtga	
gctcggtgg gtcgtgagc	
ccgtgtctgc cgagggccca	
gtcggcgacg tgggggttgg	
cgctgaggaa ggaagtccag	
agatccacgg ccagggcggt	
ctgcaagcgg tcccggtact	
gacggaaactg ctggcccacg	
gccataaaaa cgggggttgc	
gcagttagaaag gtgcgggggt	
cgccgtgcac ggggtccac	
ttgagcttgg gggcgagg	
gtgggcgagc tcgacgagcg	
gccccgtcccc ggagatttc	
atgaccagca tgaaggggac	
gagctgtttt ccgaaggacc	
ccatccaggt ttaggtttcc	
acatcgtagg tgaggaagag	
ccttcggc cgaggatgcg	
agccgatggg gaagaactgg	
atctcctgcc accagtttgg	

ggaatggctg ttgatgtgat	ggaagttagaa atgcccacgg	cgcgccgagc actcggtgtt	7980
gtgtttatac aagcgtccgc	agtgcgtcgca acgctgcacg	ggatgcacgt gctgcacgag	8040
ctgtacctgg gttcctttga	cgaggaattt cagtgggcag	tggagcgtcg gcggtgtcat	8100
ctgggtgtgt actacgttcct	ggccatcgcc gtggccatcg	tctgcctcga tgggggtcat	8160
gctgacgagc cccgcgggaa	ggcagggtcca gacttcggct	cggacgggtc ggagagcggag	8220
gacgagggcg cgccaggccgg	agctgtccag ggtcctgaga	cgctgcggag tcaggtcagt	8280
gggcagcggc ggccgcgggt	tgacttgcag gagctttcc	agggcgcgcg ggaggtccag	8340
atggtaacttgc	atctccacgg cggcggttgtt	ggcgcacgtcc acggcttgc	8400
ccccctggggc	ggccaccaccc	ggccatcgcc tgccgggtcat	8460
aagcggccgc	gaggacgcgc	ggccggccgc agggccggct	8520
ggcaggggca	cgtcgccgc	ggccgcgggc aggttctgggt	8580
gcgtgagcga	cgacgcgcacg	gttgcgtcc tggatctgac	8640
ggaccctgtga	gtttgaacct	gcctctgggt actgcgcggc	8700
gcggcctgc	gcaggatctc	ttgcacgtcg cccgagttgt	8760
atgaactgt	cgatctctc	cctggtaggc gatctcggtc	8820
gcgagggttgt	tccctgaagg	tctccgcggc cggcgccgtc	8880
cagacgcggc	tgttagaccac	gacggatgtcc	8940
aggttgcgt	cgacgtggcg	cgtaagacc gcgtagttgc	9000
ttgagcgtgg	tggcgatgtg	agaggcgctg gtagaggttag	9060
atctcgctga	cgtcgcccaag	ttgacgtccagc	9120
aagttgaaaa	actgggagtt	ggccgcgcgac	9180
agctcagcga	tggggcgcg	acggtaact cctctccag	9240
atctcttcct	cctccactaa	aagacggatg	9300
ggggccctgc	catctctct	agctcagcga tggatcttcc	9360
cctcgccggc	acttcctct	cgaggaggcg	9420
aagacgcgc	gtcgccggcg	cgccacggc	9480
cgccatctc	gacggatgtt	agacgggtcg	9540
caggtggccg	ctcggtgacg	cctcgccggg	9600
ccgggggggt	gacggatgtt	ccgcagcgtg	9660
ctccgttggg	caattggccc	cgacgttgg	
cgatcttat	gtaggactc	cgccatctc	
caattggccc	cgccatctc	cctcgacgtc	
gtaggactc	aaaccgcgtga	cgacgttgg	
cgatcttat	acgaaggctt	cgccatctc	
aaaccgcgtga	cgagccagtc	cgccatctc	
acgaaggctt	gcagtcgcaa	cgccatctc	
cgagccagtc	ggtaggctga	cgccatctc	
gcagtcgcaa	gcccggtttc	ttgttcttcg	
ggtaggctga	gggatttcgg	gaggcgccgg	
gcccggtttc	ggtaggctga	ggcgatgtcg	

ctggatga agttgaagta ggcggtcctg agacggcgg a tggtggcgag gaggcaccagg 9720
 tcctggcc cggcttgctg gatgcgcaga cggcggcca tggcccaggc gtggtcctga 9780
 cacctggcga ggtccttcta gtatgcctgc atgagccgtt ccacgggac ac ctccttcctcg 9840
 cccgcgcggc cgtgcattgcg cgtgagcccc aaccgcgtt ggggctggac gagcgcagg 9900
 tcggcga cgcgcgcggc gaggatggcc tgctgtatct gggtaggggt ggtcttggaa 9960
 tcgtcgaagt cgacgaagcg gtggtaggtt cccgtgttga tggtatagga gcagttggcc 10020
 atgacggacc agttgacggt ctggtggccg ggtcgacga gctcggttga cttgaggcgc 10080
 gagtaggcgc gcgtgtcgaa gatgtatgc ttgcagggtgc gcacgaggta ctggatccg 10140
 acgaggaagt gcggcggcgg ctggcggttag agcggccatc gctcggttgc gggggcgcc 10200
 ggcgcgaggt cctcgagcat gaggcggtgg tagccgtaga tgtaccttga catccaggtg 10260
 atgccggcgg cggcggttga ggcgcgcggg aactcgccga cgcgggttcca gatgttgcgc 10320
 agcggcagga agtagttcat ggtggcccg gcttggcccg tgaggcgccgc gcagtcgtgg 10380
 atgtctttaga catacggca aaaacgaaaag cggtcagcgg ctcgactccg tggcctggag 10440
 gctaaagcga cgggttgggc tgcgcgttga ccccggttgc aatctcgaat caggctggag 10500
 cccgagctaa cgtggtaactg gcactccgt ctcgacccaa gcctgttaac gaaaccttca 10560
 ggatacggag gcggttgcgtt ttttggcctt ggtcgcttgc catgaaaaac tagtaagcgc 10620
 ggaaagcgcac cggccgcgtat ggctcgctgc cgtatcttgc agaaagaatc gccagggttgc 10680
 cgttgcgttgc tggcccggtt cggccctcag cgcgcgttgc cggccggatt ccgcggctaa 10740
 cgtggcggtg gcttggccgt cgtttccaa accccttagc cagccgactt ctccagttac 10800
 ggagcgcgc ccttttttgc ttgttttttgc cccagatgc tcccgacttgc cggcagatgc 10860
 gcccccaccc tccacctcaa cggccctac cggccgcgtat cccgcgttcca 10920
 tggcccccgc ccagcagcag ccagccacta cggccggcggc cgcgttgcggc ggagccggcg 10980
 ttcagttatgc cctggcccttgc gaagagggcg agggcttgc cggcgttgggg gcttcgttgc 11040
 cggagcggca cccgcgcgttgc cagatgaaaa gggacgttgc cggccgttac gtgccaagc 11100
 agaacctgtt cagagacagg agcggcgttgc agcccgagga gatgcgcgc tcccgcttcc 11160
 acgcggggcg ggagcttgcgg cggcccttgc accgaaagcg ggtgttgcgg gacgaggatt 11220
 tcgaggcgga cggatcgttgcg gggatcgttgc cccgcgcgtat gcacgttgcggc gggccaaacc 11280
 tggtcacggc gtacgagcag accgttgcagg aggagagcaa cttccaaaaaa tccttcaaca 11340

accacgtgcg cacgctgate gcgcgcgagg aggtgaccct gggcctgatg cacctgtggg 11400
 acctgctgga ggccatcgta cagaacccca cgagcaagcc gctgacggcg cagcttttc 11460
 tggtggtgca gcacagtgg gacaacgaga cgttcaggga ggcgctgctg aatatcaccg 11520
 agcccgggg ccgctggctc ctggacctgg tgaacattct gcagagcatc gtggtgcagg 11580
 agcgcgggct gccgctgtcc gagaagctgg cggctatcaa cttctcggtg ctgaggctgg 11640
 gcaagtacta cgcttaggaag atctacaaga ccccgtacgt gccccatagac aaggaggtga 11700
 agatcgacgg gttttacatg cgcatgaccc taaaagtgtct gaccctgagc gacgatctgg 11760
 gggtgtaccc caacgacagg atgcaccgcg cggtagcgc cagccgcccgg cgcgagctga 11820
 gcgaccagga gctgatgcac agcctgcagc gggccctgac cggggccggg accgaggggg 11880
 agagctactt tgacatgggc gcggacctgc gctggcagcc cagccgcccgg gccttggaaag 11940
 ctgccggcgg ttccccctac gtggaggagg tggacgatga ggaggaggag ggcgagtacc 12000
 tggaaagactg atggcgcgac cgtatTTTG cttagatgcag caacagccac cgcctccctga 12060
 tccccgcgtg cggcggcgc tgcagagcca gccgtccggc attaactctt cggacgattt 12120
 gacccaggcc atgcaacgc tcatggcgct gacgaccgc aatcccgaag ccttttagaca 12180
 gcagecctcag gccaaccggc tctcgccat cctggaggcc gtggtgccct cgcgctcgaa 12240
 cccccacgcac gagaaggtgc tggccatgt gaacgcgtg gtggagaaca aggccatccg 12300
 cggcgacgag gcccggctgg tgtacaacgc gctgctggag cgcgtggccc gtcacaacag 12360
 caccaacgtg cagacgaacc tggaccgc tggaccgc gtcgcgcagg cgggtgcga 12420
 gcgcgagcgg ttccacccgc agtcgaacct gggctccatg gtggcgctga acgccttcct 12480
 gagcacgcag cccgccaacg tgccccgggg ccaggaggac tacaccaact tcatcagcgc 12540
 gctgcggctg atggtggccg aggtgcccc gacgcgggtg taccagtcgg ggccggacta 12600
 ctctttccag accagtcgccc agggcttgc gaccgtgaac ctgagccagg ctttcaagaa 12660
 cttgcaggga ctgtggggcg tgcaggcccc ggtcgccggac cgcgcgacgg tgctcgacct 12720
 gctgacgccc aactcgcgc tgcgtctgt gctggtgccg cccttacgg acagcggcag 12780
 cgtgacgccc gactcgtacc tgggtaccc gtcggcttgc taccgcgagg ccatcgcc 12840
 ggcgcacgtg gacgagcaga ctcaccagga gatcaccac gtgagccgcg cgctggcc 12900
 ggaggacccg ggcaacctgg aggcacccct gaaaccttgc ctgaccaacc ggtcgacgaa 12960
 gatccccccc cagtacgcgc tgcggccgaa ggaggagcgc atcctgcgt acgtgcagca 13020
 gagcgtgggg ctgttcctga tgcaggagggg ggccacgccc agcgccgcgc tcgacatgac 13080

cgcgcgcaac atggagccca gcatgtacgc tcgcaaccgc ccgttcatca ataagctgat 13140
ggactacttg catcgcccgg cccatgaa ctggactac ttaccaacg ccattttgaa 13200
ccccactgg ctcccgccgc cccgggttcta cacggcgag tacgacatgc cccacccaa 13260
cgacgggttc ctgtggacag acgtggacag cagcgtttc tcgcccggcc cccaccac 13320
cgtgttgaag aaagaggcg gggacggcg gccgtctcg gcgctgtccg gtcgcgcggg 13380
tgctgcccg gcgggtcctcg aggcccccag cccctcccg aecctgcct tttcgctgaa 13440
cagcgtgcgc agcagcgcgac tgggtcggt gacggcgccg cgcctgtgg gcgaggagga 13500
gtacctgaac gactccttgt tgaggcccga gcgcgagaag aacttccca ataacggat 13560
agagagcctg gtggacaaga tgagccgtg gaagacgtac ggcacgcgac acagggacga 13620
gccccagct agcagcagcg caggcacccg tagacgcccag cgacacgaca ggcagcgggg 13680
tctggtgtgg gacgatgagg attccggcgta cgacagcgcg acgttggact tgggtggag 13740
tggtgtgtgt aacccgttcg ctcacttgtc ccccgatc gggcgctga tgtaagaatc 13800
tgaaaaaata aaaaacggta ctcaccaagg ccatggcgac cagcgtgcgt tcttcgt 13860
tgttttagt agtatgatga ggccgcgtgtc cccggagggt cctcctccct cgtacgagag 13920
cgtgatgcag caggcggtgg cggcgccgt gcaagcccccgttggaggcgcttacgtgcc 13980
cccgccgtac ctggcgccata cggagggcg gaacagcatt cgtaactcg agctggcacc 14040
cttgcgtat accacccgtt tgtacctggt ggacaacaag tcggcggaca tcgcctcgct 14100
gaactaccag aacgaccaca gcaacttcct gaccaccgtgtcagaaca acgatttcac 14160
ccccacggag gccagcaccc agaccatcaa ctttgcgtgg cgctcgccgt gggcgccca 14220
gctaaaaacc atcatgcaca ccaacatgcc caacgtgaac gagttcatgt acagcaacaa 14280
gttcaaggcg cgggtgatgg tctcgccaa gaccccaat ggggtcgccgg tggatgagaa 14340
ttatgatggt agtcaggacg agctgactta cgagtgggtg gagtttgagc tgcccgggg 14400
caacttcgtgtgaccatgac ccattcgatct gatgaacaac gccatcatcg acaactactt 14460
ggcgggtgggg cgtcagaacg ggggtcgga gagcgacatc ggcgtgaagt tcgacacgcg 14520
caacttcgg ctgggctggg accccgatcg cagactggtg atgcggcgcc tgcacaccaa 14580
cgaggccttc caccccgaca tcgtcctgtc gcccggctgc ggcgtggact tcaccgag 14640
ccgcctcagc aacctgctgg gcatccgcaa gggcagccc ttccaggagg gcttccagat 14700
cctgtacgag gacctggagg gggcaacat ccccgccgtc ttggatgtcg aagcctatga 14760

gaaaagcaag gaggaggccg ccgcagcggc gaccgcagcc gtggccaccc cctctaccga 14820
 ggtgcggggc gataattttg ctagcgcgcg ggcagtggcc gaggcggtg aaaccgaaag 14880
 taagatagtc atccagccgg tggagaagga cagcaaggac aggagctaca acgtgctcgc 14940
 ggacaagaaa aacaccgcct accgcagctg gtacctggcc tacaactacg gcgaccccga 15000
 gaagggcgtg cgctccttga cgctgctcac cacctcgac gtcacctgctg gcgtggagca 15060
 agtctactgg tcgctgccc acatgatgca agacccggtc accttccgct ccacgcgtca 15120
 agtttagcaac taccgggtgg tgggcggcga gctctgccc gtctactcca agagtttctt 15180
 caacgagcag gcggctact cgcagcagct ggcgccttc acctcgctca cgcaacgttctt 15240
 caaccgccttc cccgagaacc agatcctcgt cccggccccc ggcgcacca ttaccaccgt 15300
 cagtgaaaac gttcctgctc tcacagatca cgggaccctg cgcgtgcga gcgtatccg 15360
 gggagtccag cgcgtgaccg tcactgacgc cagacgccgc acctgcctt acgtctacaa 15420
 ggcctgggc gtagtcgcgc cgcgcgtcct ctcgagccgc accttctaaa aaatgtccat 15480
 tctcatctcg cccagtaata acaccggttg gggctgcgc ggcgcacca agatgtacgg 15540
 aggcgctcgc caacgctcca cgcaacaccc cgtgcgcgtg cgcgggcaact tccgcgtcc 15600
 ctggggcgcc ctcaaggggcc gcgtgcgtc ggcgcaccc gtcgacgcacg tgatcgacca 15660
 ggtggtgccc gacgcgcga actacacgc cgcgcgcgc cccgcctcca cgcgtggacgc 15720
 cgtcatcgac agcgtggtgg ccgatgcgc cggtaacgc cgcgcacca ggcggccggc 15780
 ggcgcaccc cggccggcacc ggagcaccccc cgcgcgcgc ggcggcgag cttgtctgcg 15840
 cagggccagg cgcacgggac gcaggccat gtcaggcg gccagacgc cggccctccgg 15900
 cagcagcagc gcccgcagga cccgcacgc cgcgcacgc ggcggccgg cggccatcgc 15960
 cagcatgtcc cgcgcgcgc gcgcaacgt gtactgggt cgcgcacgc caccgggtgt 16020
 ggcgcgtgccc gtgcgcaccc gccccctcg cacttgaaga tgctgacttc gcatgtgtga 16080
 tgtgtcccag cggcgaggag gatgtccaag cgcacccata aggaagagat gtcgcggc 16140
 atcgccgcctg agatctacgg ccccgccgtg aaggaggaaa gaaagccccg caaactgaag 16200
 cgggtcaaaa aggacaaaaaa ggaggaggaa gatgtggacg gactgggtgg atttgtgcgc 16260
 gagttcgccc cccggccggc cgtgcagtgg cgcggccgg aagtgaaacc ggtgctgcgg 16320
 cccggccacca cgggtggcctt caccgcgc gacgcgtccg gtcgcctc caacgcgtcc 16380
 tacgacgagg tgtacggggc cgaggacatc ctcgagcagg cggtcgacgc tctggccgg 16440
 tttgtttacg gcaagcgac cggccggcgc cccttggaaag aggaggcggt gtccatcccg 16500

ctggaccacg gcaaccccac gccgagcctg aagccggta ccctgcagca ggtgctgccg 16560
agcgccggcgc cgccgggggg cttcaagcgc gagggccggcg aggtatctgta cccgaccatg 16620
cagctgatgg tgcccaagcg ccagaagctg gaggacgtgc tggagcacat gaaggtggac 16680
cccgaggatgc agcccgaggt caaggtgcgg cccatcaagc aggtggcccc gggcctggc 16740
gtgcagacccg tggacatcaa gatccccacg gagcccatgg aaacgcagac cgagcccgtg 16800
aagcccagca ccagcaccat ggaggtgcag acggatccct ghatgccggc gccggcttcc 16860
accactcgcc gaagacgcaa gtacggcgcg gccagcctgc tgatccccaa ctacgcgctg 16920
catccttcca tcattccccac gccggctac cgccgcacgc gettctaccc cggtacacc 16980
agcagccgccc gcaagaccac cacccggcgc cgccgtcgcc gcacccgcg cagcagcacc 17040
gcgacttccg cccggccctt ggtgcggaga gtgtacccgca gccccggcga gcctctgacc 17100
ctgcccgcgc cgccgtacca cccgagcatc gccatthaac tctgcccgtc cctcttactt 17160
gcagatatgg ccctcacatg cccgcctccgc gtcctccat cgggttacccg aggaagaaaag 17220
ccgcgcgcgtt gaaggctgac gggaaacggg ctgcgtcgcc atcaccaccc gccccggcgc 17280
gccatcagca agcggttggg gggaggcttc ctgcccgcgc tgatccccat catcgccgcg 17340
gcatcgccccgg ctagcttccgc gatccggatcc gatccgtgc aggcttccat ggcacttgc 17400
gacacagctt gaaaaatttg taataaaaaaa atggactgac gtcctggtc ctgtatgtg 17460
tgtttttaga tggaaagacat caattttcg tccctggcac cgccacacgg cacgcggccg 17520
tttatggca cctggagcga catcgccaaac agccaaactga acggggggcgc cttcaattgg 17580
agcagtctct ggagcgggct taagaatttc gggccacgc tcaaaaccta tggcaacaag 17640
gcgtggaaca gcagcacagg gcagggcgtg agggaaaagc tgaaagagca gaacttccag 17700
cagaaggatgg tcgatggcct ggcctcgggc atcaacgggg tggatggaccc gccaaccag 17760
gccgtgcaga aacagatcaa cagccgcctg gacgggtcc cgcccgccgg gtccgtggag 17820
atggcccaagg tggaggagga gtcgcctccc ctggacaagc gccccggcgc gtcgtggatgg 17880
cccgacgcgg aggagacgct gtcgtgcac acggacgcgc cgcccccgtt cgaggaggcg 17940
gtgaaaactgg gtctgcccac cacgcggccc gtggcgctc tggccaccgg ggtgctgaaa 18000
cccagcagca gcagccagcc cgccgcacccgt gacttgcctc cgccgtctc ccccccctcc 18060
acagttggcta agccccctgcc gcccggcc gtcgcgtcgcc gcccccccg aggccgcggc 18120
caggcgaact ggcagagcac tctgaacacgc atcgtggtc tggagtgca gagtgtgaag 18180

cgccggcgt gctattaaaa gacactgtag cgcttaactt gcttgtctgt gtgtatatgt 18240
atgtccgccc accagaagga ggaagaggcg cgtcgcccag ttgcaagatg gccaccccat 18300
cgatgtgcc ccagtggcg tacatgcaca tcgccccaca ggacgcttcg gagtacctga 18360
gtccgggtct ggtgcagttc gcccgcgcca cagacaccta cttcagtctg gggaaacaagt 18420
ttaggaaccc cacggtggcg cccacgcacg atgtgaccac cgaccgcagc cagcggctga 18480
cgctgcgttt cgtgcggcgtg gacggcgagg acaacaccta ctcgtacaaa gtgcgtaca 18540
cgctggccgt gggcgacaac cgcgtgtgg acatggccag cacctacttt gacatccgctg 18600
gcgtgctgga tcggggggcc agcttcaaacc cctactccgg caccgcctac aacagcctgg 18660
ctccccaaagg agcgcccaac acttgccagt ggacatataa agctggtgat actgatacag 18720
aaaaaaaaacta tacatatgga aatgcacactg tgcaaggcat tagcattaca aaggatggta 18780
ttcaacttgg aactgacagc gatggtcagg caatctatgc agacgaaact tatcaaccag 18840
agcctcaagt gggtgatgct gaatggcatg acatcaactgg tactgatgaa aaatatggag 18900
gcagagctct taagcctgac accaaaatga agccttgcta tggttcttt gccaagccta 18960
ccaataaaaga aggaggccag gcaaatgtga aaaccgaaac aggccgtacc aaagaatatg 19020
acattgacat ggcatttttc gataatcgaa gtgcagctgc cgccggccta gccccagaaa 19080
ttgtttgtt tactgagaat gtggatctgg aaactccaga tacccatatt gtatacaagg 19140
caggtacaga tgacagtagc tcttctatca atttgggtca gcagtccatg cccaaacagac 19200
ccaaactacat tggcttcaga gacaactta tcggtctgtat gtactacaac agcaactggca 19260
atatgggtgt actggctgga caggcctccc agctgaatgc tgtggtgac ttgcaggaca 19320
gaaacaccga actgtcctac cagctttgc ttgactctct gggtgacaga accaggtatt 19380
tcagtatgtg gaatcaggcg gtggacagtt atgacccga tgtgcgcatt attgaaaatc 19440
acgggtgtgga ggatgaactt cctaactatt gcttccctt ggatgctgtg ggtagaactg 19500
atacttacca gggatataag gccaatggtg ataatcaaacc cacctggacc aaagatgata 19560
ctgttaatga tgctaatgaa ttggcaagg gcaatcctt cggccatggag atcaacatcc 19620
aggccaaacct gtggcggAAC ttccctctacg cgaacgtggc gctgtacctg cccgactcct 19680
acaagtacac gcccggccaaac atcacgtgc ccaccaacac caacacctac gattacatga 19740
acggccgcgt ggtggcgccc tcgctggtg acgcctacat caacatgggg ggcgcgtgg 19800
cgctggaccc catggacaac gtcaacccct tcaaccacca ccgcaacgcg ggcctgcgt 19860
accgcctccat gtcctgggc aacggggcgct acgtgcctt ccacatccag gtgcggccaaa 19920

agttttcgc catcaagagc ctccctgtcc tgccccggtc ctacacctac gagtggaaact 19980
 tccgcaagga cgtcaacatg atccctgcaga gctccctcggt caacgacctg cgcacggacg 20040
 gggccatccat cgccttcacc agcatcaacc tctacgccac cttcttcccc atggcgcaca 20100
 acaccgcctc cacgctcgag .gccatgctgc gcaacgacac caacgaccag tccttcaacg 20160
 actacacctc ggcggccaaac atgctctacc ccattcccgcc caacgcccacc aacgtgccc 20220
 tctccatccc ctgcgcacaac tggcccgctt tccgcggctg gtccctcact cgccctcaaga 20280
 cccgcgagac gccctcgctc ggctccgggt tcgaccccta ctgcgtctac tcgggctcca 20340
 tcccctacct cgacggcacc ttctacacta accacacctt caagaaggta tccatcacct 20400
 tcgactcctc cgtcagctgg cccggcaacg accgcctctt gacgcccac gagttcgaaa 20460
 tcaagcgcac cgtcgcacgga gaggggtaca acgtggccca gtgcaacatg accaaggact 20520
 ggttccctgggt ccagatgctg gcccactaca acatggctta ccagggcttc tacgtgccc 20580
 agggtaccaa ggaccgcata tactccttct tccgcaactt ccagcccatg agccgccagg 20640
 tcgtggacga ggtcaactac aaggactacc aggcgcgtac cctggccatc cagcacaaca 20700
 actcgggctt cgtcggctac ctgcgccca ccatgcgcca gggccagccc taccggcc 20760
 actaccccta cccgctcatc ggcaagagcg ccgtcgccag cgtaaccctg aaaaagtcc 20820
 tctcgacccg ggtcatgtgg cgcatccct tctccagcaa ctgcgttcc atggcgccg 20880
 tcacccgaccc cggccagaac atgctctacg ccaactccgc ccacgcgcta gacatgaatt 20940
 tcgaagtcga ccccatggat gagtccaccc ttctctatgt tgtcttcgaa gtcttcgacg 21000
 tcgtcccgagt gcaccagccc caccgcggcg tcatcgagcc cgtaaccctg cgacgcctc 21060
 tctcgccgg caacgcacc acctaagctt ctgcgttcc tcaagatgac ggcctgcg 21120
 ggctccggcg agcaggagct cagggccatc ctccgcgacc tgggctgcgg gccccttc 21180
 ctgggcaccc tcgacaagcg ctccccggta ttcatggccc cgcacaagct ggcctgcg 21240
 atcgtaaca cggccggcccg cgagaccggg ggcgagact ggctggcctt cgcctggaaac 21300
 cccgcctccc acacctgcta cctcttcgac cccttcgggt tctcgacga ggcctcaag 21360
 cagatctacc agttcgagta cgagggcctg ctgcgtcgca ggccttcggc caccgaggac 21420
 cgctgcgtca ccctggaaaa gtccacccag accgtgcagg gtccgcgtc ggcgcctgc 21480
 gggctttctt gctgcgtt cctgcacgccc ttgcgtcact ggcccgaccg cccatggac 21540
 aagaacccca ccatgaactt gctgacgggg gtcggcaacg gcatgctcca gtcgcggcc 21600

gtggAACCCA ccctgcGCCG caaccaggAG gCGCTCTACC gCTTCCTCAA CGCCCACTCC 21660
gcctactttc gCTCCCACCG CGCGCGCATC gagaaggCCA CGCCTTCGA CGCATGAAT 21720
caagacatgt aatccggtgt gtgtatgtga atgctttatt catcataata aacagcacat 21780
gtttatGCCA CCTTCTCTGA ggctctgact ttatTTAGAA atcgaaggGGG ttctGCCGGC 21840
tctcgGCATG GCCCAGGGC agggatacgt tgCGGAACTG gtacttGGGC AGCCACTTGA 21900
actcggggat cAGCAGCTC ggcACGGGGa ggtcGGGGaa CGAGTCGCTC cacAGCTTGC 21960
gCGTgAGTTG CAGGGCGCC AGCAGGTcGG GCGCGGAGAT CTTGAAATCG CAGTTGGAC 22020
CCGCGTTCTG CGCGCGAGAG ttacggTACA CGGGGTTGCA GCACTGGAAC ACCATCAGGG 22080
CCGGGTGCTT CACGCTCGCC AGCACCGTCG CGTCGGTgAT GCCCTCCACG TCCAGATCCT 22140
CGGCGTTGGC CATCCCAGG 9gggtcatct tgCAGGTCTG CGGCCCCATG CTGGGCACGC 22200
AGCCGGGCTT GTGGTTGCAA TCGCAGTGCA GGGGGATCAG CATCATCTGG GCCTGCTCG 22260
AGCTCATGCC CGGGTACATG GCCTTCATGA AAGCCTCCAG CTGGCGGAAG GCCTGCTCG 22320
CCTTGCCTGC CTCGGTGAAG AAGACCCCGC AGGACTTGT AGAGAACTGG TTGGTGGCGC 22380
AGCCAGCGTC GTGCACGCAg CAGCGCGGT CGTTGTTGGC CAGCTGCACC ACGCTGCGCC 22440
CCCAAGCGGTT CTGGGTGATC TTGGCCCCGT CGGGGGTCTC CTTAGCGCG CGCTGCCCGT 22500
TCTCGCTCGC CACATCCATC TCGATCGTGT GTCCTTCTG GATCATCACG GTCCCGTGCA 22560
GGCACCCAG CTTGCCCTCG GCCTCGGTGC ACCCGTGCAG CCACAGCGCG CAGCCGGTGC 22620
TCTCCCAgTT CTTGTGGCG ATCTGGAGT GCGAGTGCAC GAAGCCCTGC AGGAAGCGGC 22680
CCATCATCGT GGTCAGGGTC TTGTTGCTGG TGAAGGTCAg CGGAATGCCG CGGTGCTCT 22740
CGTTCACATA CAGGTGGCAG ATACGGCGGT ACACCTCGCC CTGCTCGGGC ATCAGCTGGA 22800
AGGCAGGACTT CAGGTGCGTC TCCACCGGT ACCGGTCCAT CAGCAGCGTC ATCACTTCCA 22860
TGCCTTCTC CCAGGCCGAA ACGATCGGCA GGCTCAGGGG GTTCTTCACC GTTGTATCT 22920
TAGTCGCCGC CGCCGAAGTC AGGGGGTCGT TCTCGTCCAG GGTCTCAAAC ACTCGCTTGC 22980
CGTCTTCTC GGTGATGCGC ACGGGGGGAA AGCTGAAGCC CACGGCCGCC AGCTCTCT 23040
CGGCCTGCCT TTCTCGCTCG CTGTCCTGGC TGATGTCTTG CAAAGGCACA TGCTTGGTCT 23100
TGCAGGGGTTT CTTTTGGGC GGCAGAGGGCG GCGGGGGAGA CGTGCTGGGC GAGCGCGAGT 23160
TCTCGCTCAC CACGACTATT TCTTCTCTT GGCGTCGTC CGAGACCACG CGGCGGTAGG 23220
CATGCCTCTT CTGGGGCAGA GGCAGGGAGGCg ACGGGCTCTC GCGGTTCGGC GGGCGGCTGG 23280
CAGAGCCCT TCCCGTTCG GGGGTGCGCT CCTGGCGGCg CTGCTCTGAC TGACTTCTC 23340

cgcggccggc cattgtgttc tccttagggag caagcatgga gactcagcca tcgtcgccaa 23400
 catcgccatc tgcccccgcc gcccggacg agaaccagca gcagcagaat gaaagttaa 23460
 ccgccccgccc gcccagcccc acctccgacg ccgcagcccc agacatgcaa gagatggagg 23520
 aatccatcga gattgacctg ggctacgtga cgcccgccgaa gcacgaggag gagctggcag 23580
 cgcgttttc agccccggaa gagaaccacc aagagcagcc agagcaggaa gcagagagcg 23640
 agcagaacca ggctgggctc gagcatggcg actacctgag cggggcagag gacgtgctca 23700
 tcaagcatct ggccccccaa tgcatcatcg tcaaggacgc gctgctcgac cgcggcagg 23760
 tgccccctcag cgtggcggag ctcagccgacg cctacgagcg caaccttcc tcgcccgcg 23820
 tgccccccaa gcgcgcagccc aacggcacct gcgcgcgc 23880
 cggcttcgc ggtgcccggag gcgcgcgc 23940
 tccccgttcc ctgcgcgc 24000
 ggcgcgcct acctgatatac gcctccttgg aagaggttcc caagattttc gagggtctgg 24060
 gcagcgacga gactcgggccc gcgcgcgc 24120
 acagcgccct ggtggagttg gaaggcgaca acgcgcgc 24180
 tcgagctgac ccacttcgc tacccggcgc tcaacctgc 24240
 tcatggacca ggtgctcatc aagcgccct cgcgcgc 24300
 ccgagagctc ggacgaggccc aagcccttgg tcaagcgacga gcagctggcg 24360
 gagcgagtag cacccccccag agccttggaaag agcggcgccaa gctcatgtatg gcccgggtcc 24420
 tggtgaccgt ggagctggag tgcgtcgcc gcttcttcgc cgacgccc 24480
 aggtcgagga gAACCTGCAC taccccttca gacacgggtt cgtgcgc 24540
 tctccaaacgt ggagctgacc aacctggctt cctacatggg catcctgcac gagaaccgc 24600
 tggggcagaa cgtgctgcac accacccctgc gcggggaggc cgcgcgc 24660
 actgcgtcta cctgtacctc tgccacaccc ggcagacggg catgggcgtg tggcagcagt 24720
 gcctggagga gcagaacctg aaagagctt gcaagctct gcagaagaac ctcaaggccc 24780
 tgtggaccgg gttcgacgag cgccaccaccc cgcggaccc 24840
 agcgccctgcg gctgacgctg cgcaacgggc tgcggactt tatgagccaa agcatgttgc 24900
 aaaactttcg ctcttcatc ctgcacgct ccgggatccct gcccggcacc tgcgtcgcc 24960
 tgccctcgga ctgcgtcgcc ctgcacccgc gcgagtgccc cccggccgc 25020

gctacactgtc ggcctggcc aactacactgg cttaccactc ggacgtgate gaggacgtca 25080
ggggcgaggg cctgctcgag tgccactgcc gctgcaacct ctgcacgccc caccgctccc 25140
tggccctgcaa cccccagctg ctgagcgaga cccagatcat cggcaccctc gagttgcaag 25200
gccccggcga gggcaagggg ggtctgaaac tcaccccggg gctgtggacc tcggcctact 25260
tgcgcaagtt cgtgccccag gactaccatc cttcgagat caggttctac gaggaccaat 25320
cccagccgcc caaggcccgag ctgtcgccct gcgtcatcac ccagggggcc atcctggccc 25380
aattgcaagc catccagaaa tcccgccaag aatttctgtc gaaaaaggc cacgggtct 25440
acttggaccc ccagaccgga gaggagctca accccagtt cccccaggat gccccgagga 25500
agcagcaaga agctgaaagt ggagctgccg cccggccgg aggatttggg ggaagactgg 25560
gagagcagtc aggcagagga ggaggagatg gaagactggg acagcactca ggcagaggag 25620
gacagcctgc aagacagtct ggaggaggaa gacgaggtgg aggaggciaga ggaagaagca 25680
gccgcgcgcca gaccgtcgcc tcggcgccgg gcccacagta gatggacgca gaccggcgcc 25740
tccgtccgg gtcggggtcg cggcgccgg gcccacagta gatggacgca gaccggcgcc 25800
ttcccgAACCC ccaccaccca gacggtaag aaggagccgc agggatacaa gtcctggcg 25860
gggcacaaaa acggccatcg tccctgttg caagcctgca gggcaacat tcccttcacc 25920
cggcgttacc tgctttcca ccgggggtg aacttccccca gcaacatctt gcattactac 25980
cgtcacctcc acagccccca ctactgtttc caagaagagg cagaaaccca gcagcagcag 26040
cagcagcaga aaaccagccg cagcagctag aaaatccaca gcccggccag gtggactgag 26100
gatcgccgcg aacgagccgg cgcagacccg ggagctgagg aaccggatct ttcccacccct 26160
ctatgccatc ttccagcaga gtcggggca agagcaggaa ctgaaagtca agaaccgttc 26220
tctcgctcg ctcacccgca gttgtctgtc tcacaagac gaagaccaac ttccagcgcac 26280
tctcgaggac gccgaggctc tcttcaacaa gtactgcgc ctcactctt aagagtagcc 26340
cgcccccgc cacacacgga aaaaggccgg aattacgtca ccacctgcgc cttccggcc 26400
accatcatca tgagcaaaaga gattccacg cttacatgt ggagctacca gccccagatg 26460
ggcctggccg ccggcgccgc ccaggactac tccacccgca tgaactggct cagtgcggg 26520
cccgcgttac tctcacgggt gaatgacatc cgcgcacc gaaaccagat actcctagaa 26580
cagtcaagca tcacccgcac gccccccat cacctaattc cgcgttaattt gcccggcc 26640
ctggtgttacc aggaaattcc ccagccacg acgttactac ttcccgagaga cgcggcc 26700
gaagtccagc tgactaactc aggtgtccag ctggccggcg gcccggccct gtgtcgac 26760

cgccccgctc agggtataaa gcggctggtg atccgaggca gaggcacaca gctcaacgac 26820
gaggtggtga gctttcgct gggctgcga cctgacggag tcttccaact cgccggatcg 26880
gggagatctt ctttcacgcc tcgtcaggcc gtccctgactt tggagagtcc gtcctcgac 26940
ccccgctcggtt gttggatcggtt cactctccat ttctgtggagg agttcactcc ctccgtctac 27000
ttcaaccctt tctccggctc ccccgccac taccggacg agttcatccc gaacttcgac 27060
gccatcagcg agtcgggttga cggctacgat tgaatgtccc atggtggcgc ggctgaccta 27120
gctcggttc gacacctggta ccactgcgcg cgcttccgct gcttcgctcg ggatctcgcc 27180
gagtttgctt actttgagct gcccggagg caccctcagg gcccggccca cggagtgccgg 27240
atcgctcg aaaaaaaaaaaaaaaatcgatccac ctgttcggta tcttcagccca gcttcgatcc 27300
ctggccgagc gcgagcaagg acagaccctt ctgaccctgt actgcacatcg caaccacccc 27360
ggcctgcatg aaagtctttt ttgtctgtt tgacttgagt ataataaaag ctgagatcag 27420
cgactactcc ggaactccgtt gtgttccctgc tatcaaccag tccctgttct tcaccggaa 27480
cgagaccgag ctccagctcc agtgtaagcc ccacaagaag tacctcacct ggctgttcca 27540
gggctctccg atcgccgttg tcaaccactg cgacaacgac ggagtccctgc tgagcggccc 27600
tgccaaacctt actttttccca cccgcagaag caagctccag ctcttccaaac ctttccctccc 27660
cgggacctat cagtgcgtct cgccggccctg ccatcacacc ttccacctga tcccaatacc 27720
cacagcgatcg ctccccgcta ctaacaacca aactacccac caacgcccacc gtcgcgaccc 27780
ttccctctggg tctaatatcca ctaccggagg tgagctccga ggtcgaccaa cctctggat 27840
ttactacggc ccctggagg tggtagggtt aatagcgcta ggcctagttt cgggtggct 27900
tttggcttc tgcgttat acctcccttg ctgttcgtac ttatgttgc tgcgttgc 27960
gtttaaagaaa tggggaaagat caccctagtg agctgcggtg tgctggtggc ggtggtgctt 28020
tcgattgtgg gactggggcg cgccggctgtt gtgaaggaga aggccgatcc ctgcgttgc 28080
ttcaatcccg acaaataccca gctgaggatcc cagcccgatg gcaatcggtg cgccggctg 28140
atcaagtgcg gatggaaatg cgagaacgtg agaatacgatg acaataacaa gactcgaaac 28200
aataactctcg cgtccgtgtt gcaagccggg gaccccgagt ggtacaccgt ctctgtcccc 28260
ggtgctgacg gtcctcccgacg caccgtgaat aataacttca tttttgcgcacatgtgcac 28320
acggtcatgt ggatgagcaa gcagttacgat atgtggccccc ccacgaagga gaacatcgatg 28380
gtcttctcca tcgttacag cgtgtgcacg gctgtatccatca ccgtatcgt gtgcctgac 28440

attcacatgc tcatecgctat	tcgccccaga aataatgccg	aaaaagaaaa acagccataa	28500
cacgttttt cacacacctt	tttcagacca tggcctctgt	taaatttttg cttttatgg	28560
ccagtctcat tgccgtcatt	catggaatga gtaatgagaa	aattactatt tacactggca	28620
ctaaticcacac attgaaaggt	ccagaaaaag ccacagaagt	ttcatggtat tgttatgg	28680
atgaatcaga tgtatctact	gaactctgtg gaaacaataa	caaaaaaaaaat gagagcatta	28740
ctctcatcaa gtttcaatgt	ggatctgact taacccta	taacatcaact agagactatg	28800
taggtatgta ttatggaact	acagcaggca tttcgacat	ggaattttat caagtttctg	28860
tgtctgaacc caccacgcct	agaatgacca caaccacaaa	aactacacct gttaccacta	28920
tacagctcac taccaatggc	tttcttgcca tgcttcaagt	ggctgaaaat agcaccagca	28980
ttcaacccac cccacccagt	gaggaaattc ccagatccat	gattggcatt attgttgctg	29040
tagtggtgtg catgttgate	atgccttgc	gcatggtga ctatgccttc tgctacagaa	29100
agcacagact gaacgacaag	ctggAACACT tactaagtgt	tgaattttaa ttttttagaa	29160
ccatgaagat cctaggcctt	ttagttttt ctatcattac	ctctgctcta tgcaattctg	29220
acaatgagga cgttactgtc	gttgcggat caaattatac	actaaaaggt ccagaaaaag	29280
gtatgcttc gtggatttgt	tggttcgaa ctgacgagca	acagacagaa ctttgcattg	29340
ctcaaaaagg caaaacctca	aattctaaaa tctctaatta	tcaatgcaat ggcactgact	29400
tagtattgct caatgtcacg	aaagcatatg ctggcagtta	cacctgcct ggagatgatg	29460
ccgacaatat gatTTTtac	aaagtggaaag tggttgatcc	cactactcca cggcccacca	29520
ccacaactac tcataccaca	cacacagaac aaacaccaga	ggcagcagaa gcagagttgg	29580
ccttcaggt tcacggagat	tcctttgctg tcaatacccc	tacacccgat cagcgggtgc	29640
cgggctgtc cgtcagcggc	attgtcggtg tgctttcggg	attagcagtc ataatcatct	29700
gcatgttcat ttttgcattgc	tgctatagaa ggctttacgg	acaaaaatca gaccactgc	29760
tgaacctcta tgTTTaaattt	tttccagagc catgaaggca	gttagcgctc tagtttttg	29820
ttctttgatt ggcattgttt	ttagtgtgg gttttgaaa	aatcttacca tttatgaagg	29880
tgagaatgcc actctagtgg	gcatcagtgg tcaaaatgtc	agctggctaa aataccatct	29940
agatgggtgg aaagacattt	gcgattggaa tgtcactgtg	tatacatgta atggagttaa	30000
cctcaccatt actaatgccca	cccaagatca gaatggtagg	tttaagggcc agagtttcac	30060
tagaaataat gggatgaat	cccataaacat gtttatctat	gacgtcaactg tcacagaaaa	30120
tgagactgcc accaccacac	agatgcccac tacacacagt	tctaccacta ctaccatgca	30180

aaccacacag acaaccacta catcaactca gcatatgacc accactacag cagcaaagcc 30240
aagttagtgca ggcgcctcagc cccaggcttt ggcttggaaa gctgcacaac ctgtacaac 30300
tactaggacc aatgagcaga ctactgaatt tttgtccact gtcgagagcc acaccacagc 30360
tacccctcagt gccttcctcta gcaccgc当地 tctctcctcg ctttcctcta caccaatcag 30420
tcccgc当地 actcccaccc cagctttct cccactccc ctgaagcaaa ctgaggacag 30480
cgccatgcaa tggcagatca ccctgtcat tgtgatcggg ttggtcatcc tggccgtgtt 30540
gctctactac atcttctgcc gcccgc当地 caacgc当地 cgcaaaacgg cctacaagcc 30600
catcgttatc gggcagccgg agccgcttca ggtggaaggg ggtctaaagga atcttctt 30660
ctctttaca gtatggtcat tgaactatga ttcctagaca attcttgatc actattctta 30720
tctgc当地 ccaagtctgt gccaccctcg ctctggtggc caacgc当地 ccagactgta 30780
ttggccctt cgc当地 cttcatc ac tgc当地 ctgc当地 tgc当地 tagca 30840
tagtctgc当地 gcttatcacc ttcttccagg tcaattgactg gatcttgtg cgc当地 cgc当地 30900
acctgc当地 ccaccccaag taccgc当地 accgc当地 tggc当地 gctc aggcttctt 30960
gataaggcatg cgggctctgc tacttctcgc gcttctgtg ttagtgc当地 cccgccccgt 31020
cgaccccccgg tccccactc agtcccccgagaggtccgc aaatgcaa at tccaagaacc 31080
ctggaaattc ctcaaatttgc accgc当地 accgc当地 atcagacatg ctcccttgc当地 ggc当地 catgat 31140
cattgggatc gtgaacatcc tggc当地 gc当地 cctcatctcc tttgtgatcc acccctgtt 31200
tgacttttgtt tggactcgc cagaggc当地 ctatctccg cctgaacctg acacaccacc 31260
acagcaacct caggcacacg cactaccacc accacagc当地 aggccacaat acatgccc当地 31320
attagactat gaggccgagc cacagc当地 accgc当地 catgtcccc gctatttagtt acttcaatct 31380
aaccggc当地 gatgactgac ccactggc当地 acaacaacgt caacgacctt ctccctggaca 31440
tggacggccg cgc当地 cggag cagc当地 actcg cccacttgc cattgc当地 cagc当地 caggaga 31500
gagccgtcaa ggagctc当地 gagc当地 gatcccatag cc当地 cccacca gtgcaagaaa ggc当地 ct当地 31560
gc当地 cggtaaa acaggccaaat atcttcttacg aggtc当地 accgc当地 accat cgc当地 ct当地 31620
acgagcttgc gc当地 cggc当地 cagaagttca cctgc当地 tggactcaac cccatc当地 tca 31680
tcacccagca gtc当地 cggc当地 gatccactg cttgc当地 cggactcaac cccatc当地 tca 31740
gc当地 tccacac tctgatcaag accctctgc当地 gctccglocal cctccctcc local atgactaat 31800
cacccttta tccactgaaa taaatatcat attgatgatg atttaaataa aaaataatca 31860

tttgatttga aataaaagata caatcatatt gatgatttga gttttaaaaa ataaaagaatc 31920
 acttacttga aatctgatac caggctctg tccatgtttt ctgccaacac cacctcaet 31980
 ccctcttccc agctctggta ctgcagaccc cggcgggctg caaacttcct ccacacgctg 32040
 aaggggatgt caaattccctc ctgtccctca atcttcattt tatcttctat cagatgtcca 32100
 aaaagcgcgt cggggtgat gatgacttcg accccgtcta cccctacgt gcagacaacg 32160
 caccgaccgt gcccattcatc aacccccctc tcgtctcttc agatggattc caagagaagc 32220
 ccctgggggt gctgtccctg cgactggctg accccgtcac caccaagaac gggaaatca 32280
 ccctcaagct gggagagggg gtggacctcg actcctcgaa aaaactcata tccaacacgg 32340
 ccaccaaggc cgccgccccct ctcagttttt ccaacaacac cattccctt aacatggata 32400
 cccctcttta taccaaagat ggaaaattat ccttacaagt ttctccaccc ttaaacatata 32460
 taaaatcaac cattctgaac acattagctg tagtttatgg atcaggttta ggactgagtg 32520
 gtggactgc tcttgcagta cagttggcct ctccactcac ttttgcataaa aaaggaaata 32580
 taaaattaa cctagccagt ggtccattaa cagttgatgc aagtcgactt agtatacaact 32640
 gcaaaagagg ggtcaactgtc actacccctcg gagatgcaat tgaaagcaac ataagctggc 32700
 ctaaaggtat aagatttga ggtatggca tagctgcaaa cattggcaga ggattggaaat 32760
 ttggAACAC tagtacagag actgatgtca cagatgcata cccatttcaa gttaaattgg 32820
 gtactggcct taccttgc acgtacaggcg ccattgtgc ttggacaaa gaggatgata 32880
 aacttacatt atggaccaca gccgccccct cggccaaattt caaaatatac tctgaaaaag 32940
 atgccaaact cacactttgc ttgacaaaatgt gttggaaatgtca aattctgggt actgtgactg 33000
 tattggcagt gaataatgga agtctcaacc caatcacaaa cacagtaagc actgcactcg 33060
 tctccctcaa gtttgcata agtggagttt tgctaaagcag ctccacatta gacaaagaat 33120
 attggaaactt cagaaaggaa gatgttacac ctgctgagcc ctataactat gctatacg 33180
 ttatgcctaa cataaaggcc tattttttt acacatctgc agttcaaaa agccatattt 33240
 tcagtcagttt tatttcataat ggggatgagg ccaaaccact gatgtgttatttattt 33300
 atgaaactga ggtatgcactt tgcacccata gatcactttt tcaatggaaa tgggatagta 33360
 ctaagtacac aggtgaaaca cttgctacca gtccttcac cttctccatc atcgccccaa 33420
 aatgaacact gatccaccct ctgcatgcca acccttccca cccccactctg tctatggaaa 33480
 aaactctgaa gcacaaaata aaataaaatgtt caagtgtttt attgattcaa cagtttaca 33540
 ggattcgagc agttttttt cctccacccct cccaggacat ggaatacaccc accctctccc 33600

cccgcacagc cttgaacatc tgaatgccat tggtgatgga catgcttttgcgt 33660
 tccacacagt ttcagagcga gccagtctcg ggtcggtcag ggagatgaaa ccctccggc 33720
 actccccat ctgcacactca cagctcaaca gctgaggatt gtcctcggtg gtcggatca 33780
 cggttatctg gaagaagcag aagagcggcg gtggaatca tagtccgcga acgggatcg 33840
 ccggtgtgtc cgcatcgagc cccgcagcag tcgctgcgc cggcgtccg tcaagctgt 33900
 gctcaggggg tccgggtcca gggactccct cagcatgtg cccacggccc tcagcatcag 33960
 tcgtctggtg cggcgggcgc agcagcgc cat gctggatctcg ctcaggtcgc tgcagtacgt 34020
 gcaacacagg accaccagg tttcaacag tccatagttc aacacgtcc agccgaaact 34080
 catcgccggc agatgtcac ccacgtggcc gtcgtaccag atcctcaggt aaatcaagt 34140
 gcgcctccctc cagaacacgc tgcccacgta catgatctcc ttggcatgt ggccgttcac 34200
 cacctcccg taccacatca ccctctgggtt gaacatgcag cccggatga tccgtcgaa 34260
 ccacagggcc agcaccgcgc cggccgcatt gcagcgaaga gacccgggtt cccggcaatg 34320
 gcaatggagg acccaccgt cgtacccgtg gatcatctgg gagctgaaca agtctatgtt 34380
 ggcacagcac aggcataatgc tcatgcattt cttcagcaact ctcagctccct cgggggtcaa 34440
 aaccatatcc cagggcacgg ggaactcttgc caggacagcg aacccgcag aacagggcaa 34500
 tcctcgaca taacttacat tgtcatgga cagggatcg caatcaggca gcacccgggtg 34560
 atcctccacc agagaagcgc gggctcggtt ctcctcacag cgtggtaagg gggccggccg 34620
 atacgggtga tggcgggacg cggctgatcg tttcgacgc cgtgtcatga tgcagttgt 34680
 ttccggacatt ttcgtacttg ctgtacgaga acctggtccg ggccgtgcac accgatcgcc 34740
 ggccggggtc cggcgcttg gaacgctcgg ttttggaaatt gtaaaacagc cactctctca 34800
 gaccgtgcag cagatctagg gcctcaggag ttttggaaat cccatcatgc ctgatagctc 34860
 ttttggaaat gaccacgtt gtttggccaa gacccagccaa gatgtgcaaa ttttgggg 34920
 tttcggtgac ggcgggggag ggaagaacag gaagaaccat gattttttttaatccaaac 34980
 ggtctcgag cacttcaaaa ttttggaaat gtttggaaatt gtaaaacagc cactctctca 35040
 ggtggaaaat aacagccagg ttttggaaat gtttggaaatt gtaaaacagc cactctctca 35100
 ccagcaaagc ctccacgcgc acatccagaa acaagacaat agcggaaagcg ggagggttct 35160
 ctaattccctc aatcatcatg ttacactctt gcaccatccc cagataattt tcatttttcc 35220
 agccttgaat gattcgaact agttccttagt gtttggaaatccaa gccagccatg ataaagagct 35280

cgcgagac gcccacc ggcatttta agcacacct cataattcca agatattctg 35340
 ctctggttc acctgcagca gattgacaag cggatatca aaatctctgc cgcgatccct 35400
 aagctctcc ctcagcaata actgttagta ctcttcata tcctctccga aatttttagc 35460
 cataggacca ccaggaataa gattagggca agccacagta cagataaaacc gaagtcctcc 35520
 ccagttagca ttgccaaatg caagactgct ataagcatgc tggcttagacc cggtgatatc 35580
 ttccagataa ctggacagaa aatcacccag gcaattttta agaaaatcaa caaaagaaaa 35640
 atcctccagg tgcacgttta gggctcggg aacaacgatg aagtaaatgc aagcggtgcg 35700
 ttccagcatg gttagtttc tgatctgtaa aaaacaaaaa ataaaacatt aaaccatgct 35760
 agcctggcga acaggtgggt aaatcggttct ctccagcacc aggccaggcca cgggggtctcc 35820
 ggccgaccc tcgtaaaaat tgtcgtatg attgaaaacc atcacagaga gacgttcccg 35880
 gtggccggcg tgaatgattc gacaagatga atacaccccc ggaacattgg cgtcccgcgag 35940
 tgaaaaaaaaag cgcccgagga agcaataagg cactacaatg ctcagtctca agtccagcaa 36000
 agcgatgcca tgcggatgaa gcacaaaatc ctcaggtgctg tacaaaatgt aattactccc 36060
 ctccctgcaca ggcagcgaag ccccgatcc ctccagatac acatacaaaag ctcagcgtc 36120
 catagcttac cgagcagcag cacacaacag gcgcaagagt cagagaaagg ctgagctcta 36180
 acctgtccac ccgtctctgt ctcaatatac agcccgatc tacactgacg taaaaggccaa 36240
 agtctaaaaa tacccgccaa ataatcacac acgcccagca cacgcccaga aaccggtgac 36300
 acactcaaaa aaatacgcgc acttcctcaa acgccccaaac tgccgtcatt tccgggttcc 36360
 cacgtacgt catcgaaatt cgacttcaa attccgtcga cgttaaaaa cgtcacccgc 36420
 cccgcccccta acggtcgccc gtctctcggc caatcacctt cctccctccc caaattcaaa 36480
 cagctcattt gcataattaac ggcacccaaa agtttgaggt atattattga tgatg 36535

<210> 4
 <211> 34264
 <212> DNA
 <213> simian adenovirus SV-1

<400> 4
 tccttattct ggaaacgtgc caatatgata atgagcgggg aggagcggagg cggggccggg 60
 gtgacgtgcg gtgacgtggg gtgacgcggg gtggcgcggag ggcggggcgg gagtggggag 120
 gcgcttagtt ttacgtatg cggaggagg tttataccg gaagttgggt aatttggcgc 180
 tataacttgta agtttgtgt aatttggcgc gaaaaccggg taatgaggaa gttgaggtta 240

atatgtactt	tttatgactg	ggcggaaattt	ctgctgatca	gcagtgaact	ttgggcgtg	300
acggggaggt	ttcgctacgt	ggcagtacca	cgagaaggct	caaaggtccc	atttattgt	360
ctcctcagecg	tttcgctgg	gtattnaac	gctgtcagat	catcaagagg	ccactcttga	420
gtgccggcga	gttagagttt	ctcctcgcgc	ctgcgcgcgt	gaggctggtt	cccgagatgt	480
acggtgtttt	ctgcagcggag	acggcccgga	actcagatga	gctgcttaat	acagatctgc	540
tggatgttcc	caactcgccct	gtggcttcgc	ctccgtcgct	tcatgatctt	ttcgatgtgg	600
aagtggatcc	accgcaagat	cccaacgagg	acgcggtaaa	cagtatgtt	cctgaatgtc	660
tgtttgaggc	ggctgaggag	ggttctcaca	gcagtgaaga	gagcagacgg	ggagaggaac	720
tggacttgaa	atgctacgag	aatgtctgc	cttctagcga	ttctgaaacg	gaacagacag	780
ggggagacgg	ctgtgagtcg	gcaatgaaaa	atgaacttgt	attagactgt	ccagaacatc	840
ctgggtcatgg	ctgcccgtgcc	tgtgttttc	atagaaatgc	cagcggaaat	cctgagactc	900
tatgtgtct	gtgttatctg	cgcccttacca	gcgattttgt	atacagtaag	taaagtgttt	960
tcattggcgt	acggtagggg	attcgttgaa	gtgtttgtg	acttattatg	tgtcattatt	1020
tctaggtgac	gtgtccgacg	tggaagggga	aggagataga	tcagggctg	ctaattctcc	1080
ttgcactttg	ggggctgtgg	ttccagttgg	cattttaaa	ccgagtggtg	gaggagaacg	1140
agccggagga	gaccgagaat	ctgagagccg	gcctggaccc	tccagtggaa	gacttaggtgc	1200
tgaggatgat	cctgaagagg	ggactagtgg	gggtgctagg	aaaaagcaaa	aaactgagcc	1260
tgaacctaga	aacttttga	atgagttgac	tgttaagccta	atgaatcggc	agcgtcctga	1320
gacggtgttt	tggactgagt	tggaggatga	gttcaagaag	ggggaaattaa	acctcttgt	1380
caagtatggg	ttttagcagt	tgaaaactca	ctgggtggag	ccgtgggagg	atatggaaat	1440
ggctctagac	acctttgcta	aagtggctct	gcggccggat	aaagtttaca	ctattcgccg	1500
cactgttaat	ataaaaaaaga	gtgttatgt	tatcgccat	ggagctctgg	tgcaggtgca	1560
gaccccgac	cgggtggctt	tcaattgcgg	catgcagagt	ttgggccccg	gggtgatagg	1620
tttgaatgga	gttacatttc	aaaatgtcag	gttactggt	gatgattta	atggctctgt	1680
gtttgtgact	agcaccccagc	taaccctcca	cggtgtttac	tttttaact	ttaacaatac	1740
atgtgtggag	tcatggggta	gggtgtctct	gaggggctgc	agttttcatg	gttgcggaa	1800
ggcggtggtg	ggaagaattta	aaagtgtcat	gtctgtgaag	aaatgcata	ttgaacgctg	1860
tgtgatagct	ctagcaigtag	aggggtacgg	acggatcagg	aataacgccc	catctgagaa	1920
tggatgtttt	ctttgtctga	aaggtacggc	cagcgttaag	cataatatga	tttgcggcag	1980

cggcctgtgc ccctcgca	2040
actttaacttg cgcatatgg	2100
aactgtcaca cttgcgcac	2160
cgtcacata gtgtcccact	2220
cgccgcgcac ctggccaaca	2280
tttgagcaca atatgctat	2340
gcgttgcgc	2400
gttcacctag gtgctagacg	2460
cgccgtgttt atgccttata	2520
aatgtaaactt	2580
tagtcatact aagattttgc	2640
tggaaactga ttctttccct	2700
cgagtatgtt tcaatgggg	2760
gtttgacatg tcaatggaa	2820
ac tttttaagt gataagat	2880
at gatgaaacca agtctcg	2940
tttgcgttat cctgtaaacc	3000
tgaacgttac	3060
cgaggagctg aggacggacc	3120
accacatgt gtctgcctg	3180
cgataccact atgaaatccag	3240
tgaggtgagg ggcggagcca	3300
caaagggtat aaaggggcat	3360
gaggggtgg	3420
cgccgtgttt caaaatgagc	3480
gggacgacgg acggcaatgc	3540
gtttgaggggg ggagtgtca	3600
gccccatatct gacatctcgt	3660
cttccttcct gggcaggagt	
tcgtcagaat gtagtggct	
ccaccgtgga cggacggccg	
gtcgccctg caaaattccgc	
caccctcacc tatgccac	
tgggatcattc gttggacact	
gccgcggcag ctgccgttc	
tgctgcccgt tctactgctc	
gcggcatggc ggctgat	
ttt ggactatata accaactggc	
cactgcagct gtggcg	
ggctcttgtt tcaagaagat	
gcctgaatg tgatcttgc	
tcgcctggat atcatgtcac	
gtcgccctgga cgaaactggct	
gcatcataat cccaaatgat	
ccccgataacc gcttcagaat	
cttaaaaataa agacaaacaa	
atttgtgaa aagtaaaatg	
gctttat	
ttttttttgg	
ctcggttaggc tcgggtccac	
ctgtctcggt cgtaaggac	
tttgtgtatg tttccaaaa	
cacggtagac	
atgggcttgg atgttcaag	
atcatggcat gaggccatct	
ttgggggtgg	
gataggacca ctgaagagcg	
tcatgttccg ggggtgtatt	
gtaaatcacc cagtcgtac	
agggttttgc agcgtggAAC	
tggaatatgt cttcaggag	
ctaggctaatg gccaagggtt	
gacccttagt gtaggtgttt	
acaaagcggt tgagctgg	
gggatgcattt cgggggggaga	
tgatatgcattt ctggcttgg	
atttttaggt tagctatgtt	
accaccagg tctctgcggg	
ggttcatgtt atgaaggacc	
accagcacgg tatagccagt	
gcatttgggg aacttgcatt	
gcagtttggg gggaaaggcg	
ttggaaagaaatt tagatacccc	
cttgcgtttt cctaggtttt	
ccatgcactc atccataata	
atggcaatgg gaccctggc	
ggccgcctta gcaaaca	
cgatgtttt ggaaacatca	
tagtttgct ctagagttag	
ctcatcatag gccatcttta	
caaagcgggg taggagggtg	
cccactggg gatgatagt	
tccatctggg cctggagcgt	
atttgccttc acagatctgc	
atctcccagg ccttaatttc	
cgaggggggg atcatgtcca	

cctggggggc gataaaaaac acggtttctg gcgggggggt aatgagctgg gtggaaagca	3720
agttacgcaa cagctggat ttgcccaac cggtgggacc gtagatgacc ccgatgacgg	3780
gttcagctg gtagttcaga gaggaacagc tgccgtcggt gcgcaggagg ggagctacct	3840
cattcatcat gcttctgaca tgtttatttt cactcaactaa gtttgcaag agctctccc	3900
caccaggga taagagttct tccaggctgt tgaagtgtt cagcggttc aggccgtcggt	3960
ccatgggcat ctttcaagc gactgacgaa gcaagtacag tcggtcccag agctcggtga	4020
cgtgtcttat ggaatctcga tccagcagac ttcttggttt cgggggttgg gccgactttc	4080
gctgttagggc accagccggt gggcgccag ggccgcgagg gttctgtctt tccagggtct	4140
cagcgttcgg gtgagggtgg tctcggtgac ggtgaaggga tgagccccgg gctggcgct	4200
tgcgagggtg cgcttcaggc tcatcctgtt ggtgctgaag cggcgtcggt ctccctgtga	4260
gtcgccaga tagcaacgaa gcatgaggta gtagctgagg gactcgccg cgtgtccctt	4320
ggcgccgac ttcccttgg aaacgtgctg acatttggtg cagtgcagac acttgagggc	4380
gtagagtttt gggccagga agaccgactc gggcgagtag gcgtcggtc cgcaactgagc	4440
gcagacggtc tcgcactcca ccagccacgt gagctcggtt ttagcggat caaaaaccaa	4500
gttgcctcca ttttttttga tgcgtttctt accttgcgtc tccatgagtc tgggtcccg	4560
ttccgtgaca aaaaggctgt cggtatcccc gtagaccgac ttgagggggc gatttccaa	4620
aggtgttccg aggtcttccg cgtacaggaa ctgggaccac tccgagacaa aggctcggtt	4680
ccaggctaacc acgaaggagg cgatctgcga ggggtatctg tcgtttcaa tgaggggtc	4740
cacccccc accggtgtgca gacacaggta gtccctctcc gcgtccacga aggtgattgg	4800
cttgttaagtg taggtcacgt gacccgcacc ccccaagggt gtataaaagg gggcggtccc	4860
actctccccg tcactttctt ccgcategtt gtggaccaga gccagctgtt cgggtgagta	4920
ggccctctca aaagccggca tgatttcggc gctcaagtt tcagttctca caaacgaggt	4980
ggatttgata ttcacgtgcc ccgcggcgat gctttgtat gtggaggggt ccatctgtatc	5040
agaaaaacacg atctttttat tgtcaagttt ggtggcgaaa gacccgtaga gggcggttgg	5100
aagcaacttg gcgatggagc gcagggtctg attttctcc cgatcgcccc tctcccttggc	5160
ggcgatgtt agttgcacgt actcgccggc cacgcacccgc cactcgggga acacggcggt	5220
gcgcgtcg ggcaggatgc gcacgcgcac gccgcgggtt tgcagggtga tgaggtccac	5280
gctggtggcc acctccccgc ggaggggttc gttggtccaa cacaatcgcc cccctttct	5340
ggagcagaac ggaggcaggg gatctagcaa gttggcgggc ggggggtcggtcgatgg	5400

aaatatgccg ggttagcagaaa ttttattaaa ataatcgatt tcggtgtccg tgtcttgcaa 5460
 cgcgctttcc cacttcttca ccgcaggcgc ccttcgtag ggattcaggg gcggtccccca 5520
 gggcatgggg tgggtcaggg ccgaggcgta catgccgcag atgtcgtaca cgtacagggg 5580
 ctcccctcaac accccgatgt aagtgggta acagcgcccc ccgcggatgc tggctcgcac 5640
 gtagtcgtac atctcgtgag agggagccat gagcccgctc cccaagtggg tcttgggg 5700
 ttttccggcc cggttagagga tctgcctgaa gatggcgtgg gagttggaag agatagtggg 5760
 gcgttggaaag acgttaaagt tggctccggg cagtcccacg gagtcttggaa tgaactgggc 5820
 gtaggattcc cggagcttgt ccaccaggcgc tgccggttacc agcacgtcga gagcgcagta 5880
 gtccaaacgtc tcgcggacca ggtttaggc cgtctttgt ttttctccc acagttcgcg 5940
 attgaggagg tattcctcgc ggtctttcca gtactcttcg gcgggaaatc cttttcgtc 6000
 cgctcggtaa gaacctaaaca tgtaaaattc gttcacggct ttgtatggac aacagcctt 6060
 ttctaceggc agggcgtacg cttgagcggc ctttctgaga gaggtgtggg tgagggcgaa 6120
 ggtgtccgc accatcaatt tcaggtactg atgtttaag tccgtgtcgt cgcaggcgcc 6180
 ctgttcccac agcgtgaagt cgggtcgctt tttctgcctg ggattggggaa gggcaatgt 6240
 gacgtcgtaa aagaggattt tcccgccgcg gggcatgaag ttgcgagaga tcctgaagg 6300
 tccgggcacg tccgagcggt tggtatggac ttgcgccgc aggacgatct cgtcgaagcc 6360
 gttgatgtt gggcccacga tgtaaaatgc gataaagcgc ggctgtccct tgagggccgg 6420
 cgcttttttc aactcctcgt aggtgagaca gtccggcgcg gagagaccca gctccggcc 6480
 ggcccagtcg gagagcttag ggttagccgc gaggaaagag ctccacaggt caaggctag 6540
 cagagtttc aagcggtcgc ggaactcgcg aaactttttc cccacggcca ttttctccgg 6600
 cgtcaccacg tagaaagtgc agggcggcgtc gttccagacg tcccatcgga gctctaggc 6660
 cagctcgcag gcttgcacgaa cgagggtctc ctcgcggcag acgtgcacgaa ccagcatgaa 6720
 gggtaccaac tgtttccgcg acgagccat ccatgtgttag gttctacgt cgtaggtgac 6780
 aaagagccgc tgggtgcgcg cgtggagcc gatggggaaag aagctgatct cctgccacca 6840
 gttggaggaa tgggtgttga tgggtgaaa gtggaaatgc cggccggcgc cagagcattc 6900
 gtgtatgtt ttttccgcg gaccgcacgta gtcgcagcgc tgcacgatct gtatctcctg 6960
 aatgagatgc gcttttcgcg cgcgcaccag aaaccggagg gggaaatgtgacgggggct 7020
 tgggtggggcg gcatccccctt cgcctggcg gttggagatct gctgtcgcc cttcccttc 7080

aacacggcgg tgcgctcgcc	ggcaggatg cgcacgcgcc	agccgcgatt gtgcagggtg	5100
atgagggtcca cgctggtagc	caccccccgg cggaggggct	cgttggtcca acacaatcgcc	5160
cccccttttc tggagcagaa	cggaggcagg ggatctagca	agttggcggg cggggggtcg	5220
gcgtcgatgg tgaagatacc	ggtagcagg atcttattaa	aataatcgat ttccgtgtcc	5280
gtgtcttgca acgcgttttc	ccacttcttc accgcaggg	cccttcgtt gggattcagg	5340
ggcggtcccc agggcatggg	gtgggtcagg gccgaggcgt	acatgccca gatgtcatac	5400
acgtacaggg gttccctcaa	caccccgatg taagtgggt	aacagegccc cccgcggatg	5460
ctggctcgca cgtagtcgtt	catctcgcc gagggagcca	tgaggccgtc tcccaagtgg	5520
gtcttgggg gtttttcggc	ccggtagagg atctgtctga	agatggcgtg ggagttggaa	5580
gagatggtgg ggcgttggaa	gacgttaaag ttggcccccgg	gtagtcccac ggagttttgg	5640
atgaactggg cgtaggatttcc	ccggagtttgc	tccaccaggg cggcggtcac cagcacgtcg	5700
agagcgcagt agtccaacgt	ctcgccgacc aggtttagg	ccgtcttttgc tttttctcc	5760
cacagttcgc ggttggaggag	gtattcctcg	cggtctttcc agtacttttc ggcgggaaat	5820
ccttttcgtt ccgtcggtt	agaacctaac atgtaaaatt	cgttcacccgc tttgtatggaa	5880
caacagcctt tttctaccgg	cagggcgtac	gtttgagcgg cttttcttag agaggtgtgg	5940
gtgagggcga aggtgtcccg	caccatcaact	ttcaggtact gatgtttgaa gtccgtgtcg	6000
tcgcaggcgc cctgttccca	cagcgtgaag	tcgggtgcgt ttttctgcct gggattgggg	6060
agggcgaagg tgacatcggtt	aaagagtatt	ttcccgccgc gggcatgaa gttgcgagag	6120
atcctgaagg gcccgggcac	gtccgagcgg	ttgttgcgttgc caggacgatc	6180
tcgtcgaagc cggtgtatgtt	gtgaccacag	atgtaaaaggcgatgaaagcg cggctgtccc	6240
ttgagggccg ggcgtttttt	caactcctcg	taggtgagac agtccggcga ggagagaccc	6300
agctcagccc gggcccagtc	ggagagttga	ggattagccg caaggaagga gtcctataga	6360
tccaggcaca ggagagtttgc	caagcggtcg	cgaaactcgcc ggaactttt cccacggcc	6420
attttctccg gtgtcaactac	gtaaaaggcg	ttggggccgtt gttccacac gtcccatcggt	6480
agctctaggg ccagctcgca	ggcttggcga	acgagggctt cctcgccaga gacgtgcgt	6540
accagcataa agggtaccaa	ctgtttcccg	aacgagccca tccatgtgtt ggttctacg	6600
tcgttaggtga caaagagccg	ctgggtgcgc	gcgtggagc cgatcgaaaa gaagctgatc	6660
tcctgccacc agctggagga	atgggtgttta	atgtggtggaa agtagaagtc cccgcggcgc	6720
acagagcatt cgtgctgatg	tttgtaaaag	cgaccgcagt agtgcgcagcg ctgcacgc	6780

tgtatctcct	gaacgagatg	cgcttttgc	cgcgcacca	gaaaccggag	gggaaagtgg	6840
agacgggggg	ctgggtgggc	gacatcccc	tcgccttggc	ggtgggagtc	tgcgtctgcg	6900
tcctccctct	ctgggtggac	gacgggtgggg	acgacgacgc	cccgggtgcc	gcaagtccag	6960
atctccgcca	cggagggggtg	caggcgctgc	aggaggggac	gcagctgccc	gctgtccagg	7020
gagtcgaggg	aagtgcgcgt	gaggtcggcg	ggaagcgttt	gcaagttcac	tttcagaaga	7080
ccggtaagag	cgtgagccag	gtgcagatgg	tacttgattt	ccaggggggt	gttggatgaa	7140
gcgtccacgg	cgttagaggag	tccgtgtccg	cgcggggcca	ccaccgtgcc	ccgaggaggt	7200
tttatctcac	tcgtcgaggg	cgagcgccgg	ggggtagagg	cggctctgcg	ccggggggca	7260
gcggaggcag	aggcacgttt	tcgtgaggat	tcggcagcgg	ttgatgacga	gcccggagac	7320
tgctggcgtg	ggcgacgacg	cggcggttga	ggtcctggat	gtgcgcgttc	tgcgtgaaga	7380
ccacccggccc	ccgggtcctg	aacctaaaga	gagttccaca	aatcaatgt	ctgcacatcg	7440
aacggcggcc	tgcctgagga	tctcctgcac	gtcgcggag	ttgtctgtat	aggcgatctc	7500
ggccatgaac	tgttccactt	cttcctcgcg	gaggtcaccg	tggcccgctc	gtccacgggt	7560
ggcgccagg	tcgttggaga	tgcggcgcat	gagttgagag	aaggcggttga	ggccgttctc	7620
gttccacacg	cggctgtaca	ccacgttcc	gaaggagtcg	cgcgcgtcgca	tgaccacctg	7680
ggccacgttg	agttccacgt	ggcgccgaa	gacggcgtag	tttctgaggg	gctggaagag	7740
gtagttgagc	gtggtggcga	tgtgctcgca	gacgaagaag	tacataatcc	agcgccgcag	7800
ggtcacatctcg	ttgatgtctc	cgatggcttc	gagacgctcc	atggcctctgt	agaagtcgac	7860
ggcgaagttt	aaaaattggg	agttgcgggc	ggccacccgtg	agttcttctt	gcaggaggcg	7920
gatgagatcg	gcgaccgtgt	cgcgcaccc	ctgttcgaaa	gcgcggccag	gcgcctctgc	7980
ttcttcctcc	ggctccctct	cttccagggg	ctcggttcc	tccggcagct	ctgcgacggg	8040
gacggggcgg	cgacgtcgtc	gtctgaccgg	caggcggtcc	acgaagcgct	cgatcatttc	8100
gccgcgcgg	cgacgcattgg	tctcggtgac	ggcgctccg	ttttcgcgag	gtcgcaatcc	8160
gaagacgccc	ccgcgcagag	cgccccctg	caggagggt	aagtgggttag	ggccgtcgaa	8220
caggacacg	gcgcgtacga	tgcattttat	caattgtgc	gtaggcactc	cgtgcaggaa	8280
tctgagaacg	tcgagggtcga	cgggatccga	gaacttctct	aggaaagcgt	ctatccaatc	8340
gcaatcgcaa	gtaaagctga	gaacgggtggg	tcgtgggggg	gcgttcgcgg	gcagttggga	8400
ggtgatgctg	ctgatgatgt	aattaaagta	ggcggtcttc	aggcgccgga	tggtggcgag	8460

gaggaccacg	tctttgggcc	cggcctgttg	aatgcgcagg	cgctcggcca	tgccccaggc	8520
ctcgctctga	cagcgacgca	ggtctttgta	gaagtcttgc	atcagtctct	ccaccggaaac	8580
ctctgcttct	ccccctgtctg	ccatgcgagt	cgagccgaac	ccccgcaggg	gctgcagcaa	8640
cgcttaggtcg	gccacgaccc	tttcggccag	cacggcctgt	tgaatctgct	tgagggtggc	8700
ctggaagtctg	tccagggtcca	cgaagcggtg	ataggccccc	gtgttgcgtgg	tgttaggtgca	8760
gttggccatg	acggaccagt	tgacgacttg	catgccgggt	tgggtgatct	ccgtgtactt	8820
gaggcgcgag	taggcccctgg	actcgaacac	gtagtcgttg	catgtgcgca	ccagatactg	8880
gtagccgacc	aggaagtgag	gaggcggctc	tcgggtacagg	ggccagccaa	cggtggcggg	8940
ggcgccgggg	gacagggtcg	ccagcatgag	gcgggtggtag	tggtagatgt	agcggggagag	9000
ccaggtgatg	ccggccgagg	tgggtgcggc	cctggtaat	tcgcggacgc	ggttccagat	9060
gttgcgcagg	ggaccaaagc	gctccatggt	gggcacgctc	tgcggatgt	ggcggggcgca	9120
atcttgcacg	ctcttagatgg	aaaaaaagaca	gggcggtcat	cgactccctt	ccgttagctt	9180
gggggtaaag	tgcgaagggt	gcggcggcgg	ggaacccccc	tgcgagaccg	gccggatccg	9240
ccgcctcccg	tgcgccctggc	cccgcatcca	cgacgtccgc	gccgagaccc	agccgcgacg	9300
ctccgcggca	atacggaggg	gagtcttttg	gtgttttttc	gtagatgcac	ccgggtgtgc	9360
ggcagatgcg	accccaagacg	cccactacca	ccgcccgtggc	ggcagtaaac	ctgagcggag	9420
gcgggtacag	ggaggaggaa	gagctggctt	tagacctgga	agagggagag	gggctggccc	9480
ggctgggagc	gccccatccca	gagagacacc	ctagggttca	gtctgtgagg	gacgccaggc	9540
aggctttgt	gcccgaagcag	aacctgttta	gggaccgcag	cggtcaggag	gcggaggaga	9600
tgcgcgattg	caggtttgg	gcgggcagag	agctcaggc	gggcttcgat	cgggagcggc	9660
tcctgagggc	ggaggatttc	gagcccgacg	agcgttctgg	ggtgagcccg	gccccgcgctc	9720
acgtatccgc	ggccaaacctg	gtgagcgcgt	acgagcagac	ggtgaacgcag	gagcgcaact	9780
tccaaaagag	ctttaacaat	cacgtgagga	ccctgatcgc	gagggaggag	gtgaccatcg	9840
ggctgatgca	tctgtggac	tgcgtggagg	cctacgtgca	gaacccggct	agcaaacccc	9900
tgacggccca	gctgttccctg	atcggtcagc	acagccgcga	caacgagacg	ttccgcgcacg	9960
ccatgttcaa	catcgccggag	cccgagggtc	gctggcttt	ggatctgatt	aacatcctgc	10020
agagcatcg	ggtgcaggag	aggggcctga	gtttagcgga	caaggtggcg	gccattaact	10080
attcgatgca	gagcctgggg	aagttctacg	ctcgcaagat	ctacaagagc	ccttacgtgc	10140
ccatagacaa	ggaggtgaag	atagacagct	tttacatgcg	catggcgctg	aaggtgctga	10200

cgctgagcga cgacacctggc gtgtaccgta acgacaagaat ccacaaggcg gtgagcgcca 10260
gccggccggcg ggagactgagc gacagggagc tgatgcacag cctgcagagg gcgctggcg 10320
gcccggggca cgaggagcgc gaggcttact tcgacatggg agccgatctg cagtggcg 10380
ccagcgcgcg cgcccttggag gcggcggtt atcccgacga ggaggatcg gacgatttg 10440
aggagggcagg cgagttacgag gacgaagcct gaccggcag gtgttgtttt agatgcagcg 10500
gccggcggac gggaccaccc cggtatccgc acttttggca tccatgcaga gtcaaccttc 10560
gggcgtgacc gcctccgatg actggggcgc ggccatggac cgcatcatgg cgctgaccac 10620
ccgcaaccccc gaggctttta ggcagcaacc ccaggccaac cgtttttcgg ccatttttgg 10680
agcgggtggtg cctgcgac ccaaccgcac gcacgagaaa gtccctgacta tcgtgaacgc 10740
cctggtagac agcaaggcca tccgccgtga cgaggcggtc ttgatttaca acgcttttt 10800
ggaacgcgtg ggcgcgtaca acagcactaa cgtcagacc aatctggacc gcctcaccac 10860
cgacgtgaag gaggcgctgg cgccagaagga gcggtttctg agggacagta atctgggctc 10920
tctggtggtca ctgaaegcct tcctgagctc acagccggcc aacgtgcccc gcgggcagga 10980
ggattacgtg agtttcatca gcgccttgag actgtgtggtgc tccgagggtgc cccagagcga 11040
ggtgtaccag tctggggccgg attactttt ccagacgtcc cgacagggtc tgcaaacgg 11100
gaacctgact caggccctta aaaacttgca aggcatgtgg ggggtcaagg ccccggtgg 11160
cgatcgcgcc actatctcca gtctgctgac ccccaacact cgccctgctgc tgcttttat 11220
cgcacccgtt accaacagta gcaactatcg ccgtgactcg tacctgggtc atctcatcac 11280
tctgtaccgc gaggccatcg gccaggctca gatcgacgag catacgatc aggagattac 11340
taacgtgagc cgtgcctgg gtcaggaaga taccggcagc ctggaagcca cggttgaactt 11400
tttgctaacc aaccggaggc aaaaaatacc ctcccagttc acgttaagcg ccgaggagga 11460
gaggattctg cgatacgtgc agcagtcgt gaggctgtac ttgatgcgcg agggcgccac 11520
cgcttccacg gcttttagaca tgacggctcg gaacatggaa ccgtcctttt actccgccc 11580
ccggccgttc attaaccgtc tgatggacta cttccatcgcc gccggccgcca tgaacgggg 11640
gtacttcacc aatgccatcc tgaatccgca ttggatgccc ccgtccggct tctacaccgg 11700
ggagtttgcac ctgcccgaag ccgacgacgg ctttctgtgg gacgacgtgt ccgatagcat 11760
tttcacgccc gctaatcgcc gattccagaa gaaggaggc ggagacgagc tcccccttc 11820
cagcgtggaa gcggccctcaa ggggagagag tccctttcca agtctgtttt ccgcaggtag 11880

cggtcgggta acgcgtccac ggttgccggg ggagagcgac tacctgaacg accccttgc 11940
gcgaccggct agaaaagaaaa atttcccaa taa~~cccc~~gggtg gaaagcttgg tggataaaat 12000
gaatcg~~t~~tg aagacgtacg cccaggagca gcgggagtgg gaggacagtc agccgcggcc 12060
gctggtaccg ccgcattggc gtcgc~~c~~agag agaagacccg gacgactccg cagacgatag 12120
tagcgtgtt~~g~~ gac~~c~~ctggag ggagcggagc caacccctt~~t~~ gctca~~c~~ttgc aacccaagg 12180
gcgc~~t~~cgagt cgcc~~t~~gtatt aataaaaaag acgcggaaac ttaccagagc catggccaca 12240
gcgtgtgtgc tt~~t~~tttcc~~t~~tc t~~t~~tttcc~~t~~cc tcggcgcggc agaatgagaa gagcgg~~t~~gag 12300
agtcacgccc gcggcgtatg agggcccgcc cc~~t~~ttt~~t~~ac gaaagcgtga tgggatcagc 12360
gaacgtgccc gccac~~c~~gtgg agg~~c~~gcctta cgttcc~~t~~ccc agatacctgg gac~~c~~tacgg 12420
gggcagaaac agcatccg~~t~~t actccgagct gg~~c~~ccctt~~t~~g tacgata~~c~~cca ccaagg~~t~~gt~~a~~ 12480
cctgg~~t~~ggac aacaagtcgg cggacatcgc ctcc~~t~~gaat tac~~c~~aaaacg atcacagtaa 12540
ctttctgact accgtgg~~t~~gc agaacaatga ct~~t~~caccccg acggaggcgg gcacgc~~a~~gac 12600
cattaactt~~t~~ gac~~g~~agc~~g~~ttt cccg~~t~~gggg cgg~~t~~cag~~t~~g aaaaccatcc tgcacaccaa 12660
catgccc~~a~~ac atcaac~~g~~agt t~~c~~atgtccac caacaagttc agggcta~~a~~gc tgatgg~~t~~aga 12720
aaaaagtaat g~~c~~ggaaactc ggcagcccg atacg~~g~~atgg ttcgagttt~~a~~ ccattccaga 12780
gggcaactat tccgaaacta tgactatcga t~~t~~catgaat aac~~g~~cgat~~c~~g tggacaatta 12840
cctgcaag~~t~~g gggagacaga acgggg~~t~~gct ggaaagc~~g~~at atcggcgtga aattc~~g~~atac 12900
cagaaactt~~c~~ cgactgggg~~t~~ gggatcc~~c~~gt gac~~a~~ag~~t~~g~~t~~gatg~~c~~ccag gcgtgtacac 12960
caac~~g~~agg~~t~~ct tttcacccgg acatcgt~~t~~g~~c~~ g~~t~~gc~~c~~ccggg~~t~~ tgcgg~~t~~gtgg acttca~~c~~ct 13020
gagccg~~t~~tt~~g~~ agtaac~~c~~t~~g~~t taggaatt~~g~~ aaagcgc~~c~~gc cc~~t~~ttccaa~~g~~ aggg~~c~~ttt~~c~~a 13080
aatcatgtat gaggac~~c~~tgg agggag~~g~~taa tatacccgcc ttactggac~~g~~ tgcgaag~~t~~a 13140
cgaagct~~g~~ac atacaac~~g~~cg~~c~~ ccaaagc~~g~~ga ggtagagag attcgggg~~g~~ag acac~~c~~ttt~~g~~c 13200
ggtagctccc caggac~~c~~tgg aaatagtgcc t~~t~~taactaaa gac~~a~~g~~c~~aa~~g~~ acagaag~~c~~ta 13260
caatattata aacaacac~~g~~a cggac~~c~~cc~~t~~ gtatcg~~g~~ac~~t~~ tgg~~t~~ttctgg cttacaacta 13320
cg~~g~~agac~~cc~~cc gagaagg~~g~~ag tgagatcat~~g~~ gaccata~~c~~tc accaccac~~g~~g acgtgac~~c~~tg 13380
tggctcg~~c~~ag caagtgtact ggtccctg~~c~~cc ggat~~t~~gat~~g~~t~~g~~ caagacccgg t~~c~~ac~~c~~ttcc~~g~~ 13440
cc~~c~~ctccacc caagt~~c~~ag~~c~~a acttcc~~c~~gg~~t~~ ggtggcacc gagct~~c~~gtgc c~~c~~gtccat~~g~~c 13500
caagagcttc tac~~a~~ac~~g~~ag~~c~~c agg~~c~~cg~~t~~ct~~a~~ ctgc~~c~~a~~c~~tt attc~~g~~cc~~a~~gt ccac~~c~~gc~~g~~ct 13560
tacccac~~g~~t~~g~~ t~~t~~caatcg~~c~~t ttcc~~c~~gagaa ccagatt~~c~~tg gtgc~~c~~cc~~t~~c ccg~~c~~t~~c~~ta~~c~~ 13620

cattaccacc gtcagtgaaa acgttcccgcc cctcacagat cacggaaccc tgccgctgac 13680
 cagcagtatac agtggagttc aiegelgtac catcaccgac gccagacgtc gaacctgecc 13740
 ctacgtttac aaagcgcttg gcgtgggtggc tcctaaagtt ctttctagtc gcacccctcta 13800
 aaaacatgtc catcctcatc tctcccgata acaaaccggg ctggggactg ggctccggca 13860
 agatgtacgg cgagccaaa aggccgtcca gtcagcaccc agttcgagtt cggggccact 13920
 tcccgctcc ttggggagct tacaagcgag gactctcggt tcgaacggct gttagacgata 13980
 ccatagatgc cgtgattgcc gacgcccccc ggtacaaccc cggaccggcc gctagcgccg 14040
 cctccaccgt ggattccgtg atcgacacggt tggtagccgg cgctcgccgca tatgtctcgcc 14100
 gcaagaggcg gctgcacccgg agacgtcgcc ccacccggcc catgctggca gccaggggccg 14160
 tgctgaggcg ggcccggagg gcaggcagaa gggctatgcg ccgcgtccgc gccaacgcgg 14220
 ccgcggggag ggcccggccgaa caggctggcc gccaggctgc cgctgcccgc gctagcatgg 14280
 ccagacccag gagagggaaac gtgtactggg tgcgtgattc tgtgacggga gtccgagtgc 14340
 cggtgccgac ccgacccccc cgaagttaga agatccaaagc tgcaagacg gcggtactga 14400
 gtctccctgt tgttatcagc ccaacatgag caagcgcaag tttaaagaag aactgctgca 14460
 gacgctggtg cctgagatct atggccctcc ggacgtgaag ccagacattt aaaaaaaaaaaaa 14520
 tatcaagcggt gttaaaaaaagc gggaaaagaa agaggaactc gcggtggtag acgtggccgg 14580
 agtggaaattt attaggagtt tcgccccggc acgcagggtt caatggaaag ggcggccgggt 14640
 acaacgcgtt ttgaggccgg gcacccgggt agtttttacc ccgggagagc ggtcggccgt 14700
 taggggttcc aaaaggcagt acgacgaggt gtacggcgac gaggacatat tggaacaggc 14760
 ggctcaacag atcggagaat ttgcctacgg aaagcggtcg cgctcgcaag acctggccat 14820
 cgccttagac agcggcaacc ccacccggc cctcaaaaccc gtgacgctgc agcaggtgt 14880
 tcccgtgagc gccagcacgg acagcaagag ggggattaag agagaaatgg aagatctgca 14940
 tcccaccatc caactcatgg tccctaaacg gcagaggctg gaagaggctc tggagaagat 15000
 gaaagtggac cccagcatag agccggatgt aaaagtcaaga cctattaagg aagtggccccc 15060
 cggctttggg gtgcaaacgg tggacattca aatccccgtc accaccgctt caaccggcgt 15120
 ggaagctatg gaaacgcaaa cggagacccc tgccgcatc ggtaccaggg aagtggcggt 15180
 gcaaacggag ctttgtacg aatacgcage ccctcgccgt cagaggcggtt ccgtctgtta 15240
 cggccccggcc aacgcacatca tgccagaata tgccgtgcattt ccgtcttaccc 15300

cggtatccgg ggtgtgacgt atcgcccgtc tggaacccgc cgccgaaccc gtcgccgccc 15360
 ccgcctccgt cgcgctctgg cccccgtgtc ggtgcggcgt gtgacccgccc gggaaagac 15420
 agtcgtcatt cccaaacccgc gttaccaccc tagcatecctt taataactct gccgtttgc 15480
 agatggctct gacttgccgc gtgcgecttc ccgttccgca ctatcgagga agatctcgtc 15540
 gtaggagagg catgacgggc agtggtcgccc ggcgggcttt gcgcaggcgc atgaaaggcg 15600
 gaattttacc cgccctgata cccataattt ccgcccgcatt cggtgccata cccggcggtt 15660
 cttcagtggc gttgcaagca gtcgtataata aataaaacaaa ggctttgtca cttatgacct 15720
 ggtcctgact attttatgca gaaagagcat ggaagacatc aattttacgt cgctggctcc 15780
 gcgccacggc tcgcggccgc tcatggcac ctggAACGAC atcggcacca gtcagctcaa 15840
 cggggcgct ttcaatttggg ggagcctttg gagcggcatt aaaaacttttgcgtccacgtat 15900
 taaatcctac ggcagcaaag cctggAACAG tagtgctggc cagatgctcc gagataaaact 15960
 gaaggacacc aacttccaag aaaaagtggt caatgggtg gtgaccggca tccacggcgc 16020
 ggttagatctc gccaaccaag cggcgtcagaa agagatttgc acggcggttgg aaagctcgcg 16080
 ggtgccgccc cagagagggg atgaggttggc ggtcgaggaa gtagaagttag aggaaaaagct 16140
 gcccccgctg gagaaagttc cccggcgc tccgagaccg cagaagcgcc ccaggccaga 16200
 actagaagag actctggta cggagagcaa ggagcctccc tcgtacgagc aagccttggaa 16260
 agagggcgcc tctccaccct cctaccggat gactaagccg atcgcaccca tggctcgacc 16320
 ggtgtacggc aaggattaca agcccgtaac gcttagagctg cccccacccgc cccccacggc 16380
 cccgaccgtc cccccccgtc cgactccgtc ggcggccgcg gcgggaccccg tgtccgcacc 16440
 atccgctgtc cctctgcccag ccgccccgtcc agtggccgtg gccactgcca gaaaccccg 16500
 aggccagaga ggagccaaact ggcaaaagcac gctgaacagc atcgtgggccc tgggagtgaa 16560
 aagcctgaaa cgccgcccgtt gctattatta aaaaagtgtt gctaaaaagt ctcccggtgt 16620
 atacgcctcc tatgttaccg ccagagacga gtgactgtcg ccgcgcgcgc cgctttcaag 16680
 atggccaccc catcgatgtat gccgcgtgg tcttacatgc acatcgccgg ccaggacgcc 16740
 tcggagtgacc tgagtcccg cctcggtcgag tttggccgcg ccacccgacac ctacttcage 16800
 ttgggaaaca agtttagaaa cccccccgtg gccccccaccc acgatgtgac cacggaccgc 16860
 tcgcagaggc tgaccctgcg ctttgcgtcc gtagacccggg aggacacccgc gttactttac 16920
 aaagtgcgtc acacgttggc cgtaggggac aaccgagtgcc tggacatggc cagcacctac 16980
 tttgacatcc ggggggtgtt ggtatcggggtt cccagcttca agccctatttcc cgccaccgct 17040

tacaactccc tggccccc aaaggagctccc aaccctcgga aatggacgga cacttccgac 17100
aacaaaactta aagcatatgc tcaggctccc taccagagtc aaggacttac aaaggatggt 17160
attcaggttg ggcttagttgt gacagagtca ggacaaacac cccaatatgc aaacaaagtg 17220
taccaaccccg agccacaaat tggggaaaac caatggaatt tagaacaaga agataaagcg 17280
gccccggagag tcctaaagaa agataccct atgtttccct gctatgggtc atatgccagg 17340
cccacaaacg aacaaggagg gcaggcaaaa aaccaagaag tagattaca gtttttgcc 17400
actccggcg acacccagaa cacggctaaa gtggacttt atgctaaaaa tgtcaacctg 17460
gaaactccag atactcaatt agtgtttaaa cccgatgacg acagcaccag ttcaaaactt 17520
cttcttggc agcaggctgc acctaacaga cccaaactaca taggttttag agataatttt 17580
attggtttaa tgtactacaa tagcaactgga aacatggcg tgctggccgg acaggcttct 17640
caattgaatg ccgttagtgcg ctgcaggac agaaaacaccg agttgtccta ccagctgatg 17700
ctggacgcac tgggggatcg cagccgatat tttcaatgt ggaatcaggc agtagacagc 17760
tatgacccag acgttagaat tatagaaaac cacggagtgg aagacgaact gccaaactat 17820
tgtttcctc tgggaggaat ggtggtgact gacaattaca actctgtgac gcctaaaaat 17880
ggagggcagtg gaaatacatg gcaggcagac aataactacat ttagtcaaag aggagcgcag 17940
attggctccg gaaacatgtt tgccctggaa attaacctac aggccaacct ctggcgccgc 18000
ttcttgatt ccaatattgg gttgtatctt ccagactctc tgaaaatcac ccccgacaac 18060
atcacgctgc cagaaaacaa aaacacttat cagtagatga acggtcgcgt aacgccaccc 18120
gggctcatag acacctatgt aaacgtggc gcgcgcgtggt ccccccgtgt catggacagc 18180
attaacccct tcaaccacca ccgttaacgcg ggcttgcgt accgctccat gctttggc 18240
aacggccgtt atgtgcctt tcacattcag gtgccccaaa aattcttgc cattaaaaac 18300
ctgctgcttc tccccgggttc ctataacctat gagtggact tccgcaagga tgtcaacatg 18360
atcctgcaga gctcgctggg taatgacctg cgagtggacg gggccagcat acgctttgac 18420
agcattaacc tgtatgccaa ctttttcccc atggccccaca acacggccctc taccctggaa 18480
gccatgctgc gcaacgacac caatgaccag tccttcaacg actacctgtg cgccgctaac 18540
atgctgtacc ccattcccgca acacgccacc agcgtgccc tttctattcc ttctcggaac 18600
tgggctgcct tcaggggctg gagtttact cgccctaaaa ccaaggagac tccctcgctg 18660
ggctccggtt ttgaccccta cttgtttac tccggctcca ttccctaccc agatggcacc 18720

ttttaccta accacactt caaaaaggta tctattatgt ttgactccctc ggtagctgg 18780
 cccggcaacg accgcctgtt aacgcccac gagttcgaaa ttaagcgttc cgtggacggt 18840
 gaagggtaca acgtggccca gagcaacatg accaaggact ggtttcta at tcaa at gctc 18900
 agtca ctata atataggta ccagggttc tatgtgccc agaactacaa ggaccgcatg 18960
 tactccttct tccgcaactt ccaaccaatg agccggcagg tggtagatac cgtgacttat 19020
 acagactaca aagatgtcaa gctccctac caacacaaca actcagggtt cgtggctac 19080
 atgggaccca ccatgcgaga gggacaggcc taccggcca actatccctt cccccctgatc 19140
 ggagagactg ccgtacccag cctcacgcag aaaaagtcc tctgcgaccg ggtgatgtgg 19200
 aggataccct tctctagcaa ctttatgtcg atgggctccc tcaccgaccc gggcagaac 19260
 atgctgtacg ccaactccgc tcacgcctt gacatgactt ttgagggtgga tcccatggat 19320
 gagcccacgc ttctctatgt tctgtt gaa gtctcgacg tggtagccat ccaccagccg 19380
 caccgcggcg tcatcgaggc cgtctacctg cgacacaccc tctctgccgg taacgcacc 19440
 acctaaagaa gctgatgggt tccagcgaac aggagttgca ggccattgtt cgcgacctgg 19500
 gctgcgggccc ctgcttttgc ggcacccctcg acaagcgtt tcccgattt atgtccccc 19560
 acaagccggc ctgcgcacatc gttAACACGG ccggacggga gacagggggg gtgcactggc 19620
 tcgccttcgc ctggAACCCG cgcaacccgca cctgctaccc ttgcacccct tttggttct 19680
 ccgacgaaag gctgaagcag atctaccaat tcgagtacga ggggctccctc aagcgcagcg 19740
 ctctggccctc cacgcggc acactgcgtca ccctggaaaa gtccacccag acggtccagg 19800
 ggccctccctc ggccgcctgc gggctttct gtgcacgtt tttgcacgcc ttcgtgcact 19860
 ggccctcacac ccccatggag cgcaacccca ccacggatct gtcacccgga gtgcccaca 19920
 gcatgcttca cagtcggccag gtcgcggccca ccctgcgtcg caatcaggac cacctgtatc 19980
 gctttctggg gaaacactt gcctatttcc gccgcacccg gcagcgcacatc gaacaggcca 20040
 cggcccttcga aagcatgagc caaagagtgt aatcaataaa aaccgtttt atttgacatg 20100
 atacgcgttt ctggcgttt tattaaaaat cgaagggttc gagggagggg ttcgtgcct 20160
 cgctggggag ggacacgttg cggtaactgga atcgggcgtt ccaacgaaac tcggggatca 20220
 ccagccgcgg cagggccacg tttccatgt tctgcttca aaactgtcgc accagctgca 20280
 gggctcccat cacgtcgggc gctgagatct tgaagtcgca gttagggccg gagcccccgc 20340
 ggctgttgcg gaacacgggg ttggcacact ggaacaccaa cacgctgggg ttgtggatac 20400
 tagccagggc cgtcggttcg gtcacccatcg atgcacccatcg atccctcgatc ttgctcagg 20460

cgaacggggt cagcttgcac atctgccgcc cgatctgggg taccaggatcg cgcttgtga 20520
ggcagtcgca ggcgcagaggg atgaggatgc gacgctgccc gcgttgcac atggggtaac 20580
tcgcccggcag gaactccctct atctgacgga aggccatctg ggccttgacg ccctcggtga 20640
aaaatagccc acaggacttg ctggaaaaca cgttattgcc acagttgatg tcttccgcgc 20700
agcagcgcgc atcttcgttc tttagctgaa ccacgttgcg accccagcgg ttctgaacca 20760
ccttggcttt cgtggatgc tccttcagecg cccgctgtcc gtttcgcctg gtcacatcca 20820
tttccaccac gtgctcccttg cagaccatct ccactccgtg gaaacagaac agaatgcct 20880
cctgttgggtt attgcgatgc tccccacacgg cgaccccggt ggactcccgat ctcttgtt 20940
tcaccccccgc gtaggcttcc atgtaagcca tttagaaatct gcccacatcgc tcagtgaagg 21000
tcttcgtt ggtgaaggtt agcggcaggc cgccgtgttc ctcgttcaac caagtttgcac 21060
agatcttgcg gtacacggct ccctggtcgg gcagaaaactt aaaagtctt ctgtctcg 21120
tgtccacgtg gaacttctcc atcaacatcg tcatacttc catgccccttc tcccaggcag 21180
tcaccagcgg cgccgtctcg gggttttca ccaacacggc ggtggagggg ccctcgccgg 21240
ccccgacgtc cttcatggac atttttgaa actccacggt gccgtccgcg cggcgtaactc 21300
tgccatcggtt agggtagctg aagcccacct ccatgacggt gctttcgccc tccgtgtcgg 21360
agacgatctc cggggagggc ggcggAACGG gggcagactt gcgagccctc ttcttggag 21420
ggagcggagg cacctcctgc tcgcgttcgg gactcatctc ccgcaagtag ggggtgatgg 21480
agcttcctgg ttggttctga cggttggcca ttgtatccta ggcagaaaga catggagctt 21540
atgcgcgagg aaactttaac cgccccgtcc cccgtcagcg acgaagaggtt catgtcgaa 21600
caggacccgg gctacgttac gcccggcag gatctggagg ggccttaga cgaccggcgc 21660
gacgcttagtg agcggcagga aaatgagaaa gaggaggagg agggctgcta cctcctggaa 21720
ggcgcacgttt tgctaaagca ttgcggcagg cagagcacca tactcaagga ggccttgcaa 21780
gaccgcgtccg aggtgcctt ggacgtcgcc ggcgtctccc aggcctacga ggcgaacctt 21840
ttctcgcccc gagtgccctcc gaagagacag cccaaacggca cctgcgagcc caacccgcga 21900
ctcaacttct accccgtgtt cgcgtcccc gaggcgttgg ccacctacca catcttttc 21960
aaaaaccagc gcatccccct ttccctgcccc gccaacccgc acggggccga taggaagcta 22020
acactcagaa acggagtcag catacctgat atcacgtcac tggaggaagt gcctaagatc 22080
ttcgagggtc tgggtcgaga tgagaagcgg gcccggaaacgc ctctgcagaa agaacagaaa 22140

gagagtca	acgtgctgg	tggt ggagctggag	ggggacaacg	cgcgtctgac	cgtcctcaa	22200															
cgttcat	ag aat	tttccca	cttgc	cttac	ccggccctca	acctgcc	ccc	caaagg	tat	g	22260										
aaatcggt	ca	tggacc	agc	act	catcaag	ag	agactg	agc	cc	ctgaat	cc	cgaccac	cct	22320							
gaggcgg	aaa	actc	agag	gag	ga	cg	ggaa	ag	ccc	gtcg	tca	cg	aggag	ct	cg	22380					
ctggaa	acca	ggg	aaaa	cccc	ca	g	ca	g	cc	at	tttca	g	cgac	gtg	ga	ac	gtac	gc	22440		
ctgg	tcac	gg	tgg	agc	ta	g	at	g	ct	tc	g	ca	g	at	gt	at	gt	g	22500		
aaaatcg	gggg	agt	ccct	tgca	ct	a	ca	cac	ctt	c	cc	cg	gg	ct	gc	ca	aa	22560			
atctcca	ac	g	ta	g	ct	c	a	ac	c	t	g	ttt	ca	tc	cc	ca	g	aa	22620		
ctgggg	caga	gc	gt	ct	gca	ct	g	cac	c	tt	tg	ca	gg	gg	at	cc	ga	cc	22680		
gact	gcgt	ct	ac	ctt	cc	c	c	cc	c	ttt	c	cc	ttt	cc	gt	gg	c	ag	22740		
tgctt	gga	ag	ag	aa	ac	ct	ca	aa	ag	act	cc	t	ct	g	cc	g	cc	cc	22800		
ctctgg	accg	g	ctt	cag	cg	ca	gc	ca	cgg	tc	cc	tg	ge	cc	cc	tgg	ca	cat	tttccc	22860	
gaacgc	c	ct	g	ta	g	aa	ac	cc	tt	tc	at	cg	gg	att	t	cat	c	gt	tc	22920	
caaaactt	cc	tc	ttt	cgt	c	t	ct	gg	ag	cc	tt	cc	tt	cc	at	tc	tt	cc	22980		
ctgc	ctt	ct	g	act	ttt	gt	ccc	c	ttt	c	cc	tc	cc	act	cc	gt	gg	ac	cc	23040	
tgct	ac	ct	ct	t	tc	t	c	ttt	c	cc	t	cc	tc	cc	at	tc	ttt	cc	23100		
agcgg	agagg	gg	ct	gt	ct	g	tg	cc	act	tc	g	cc	ac	t	cc	tc	cc	at	cg	23160	
ctgg	cct	gca	ac	acc	cg	aa	cc	cc	agg	gt	ca	ttt	cc	ttt	cc	at	cc	tg	at	cc	23220
gggccc	c	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	23280	
gtt	tac	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	23340	
gaccaat	ctc	g	ac	cac	cc	g	aa	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	23400	
ctgg	cc	cc	at	t	g	ca	at	cc	at	cc	at	cc	at	cc	at	cc	at	cc	at	23460	
gggg	gt	tac	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	23520	
cgag	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	23580	
agcga	aga	ag	ag	ac	tc	g	cc	ac	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	23640	
ccagg	cag	ag	gg	tt	tc	cc	ac	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	23700	
ggagg	ac	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	gg	23760	
ttt	c	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	tc	23820	
cagcc	gtt	gg	gg	gg	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	cc	23880	

caaacgcgag cggcgaaaaattgtccctg gcggacccac aaaagcagta tcgtgaactg 23940
 cttcaacac tgcggggaa acatctccctt tgcccgacgc tacctccctt tccatcacgg 24000
 tgtggcttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24060
 gctcgagaa aaaagctaag gcctcccttg ccgcgaggaa aaactccgcc gcccgtgccg 24120
 ccaaggatcc gccggccacc gaggagctga gaaagcgcattttccact ctgttatgcta 24180
 tcttcagca aagccgcggg cagcacccctc agcgcgaact gaaaataaaa aaccgctcc 24240
 tccgctcaact caccgcgacgc tgtctgtacc acaagagaga agaccagctg cagcgcaccc 24300
 tggacgacgc cgaaggactg ttcaagcaaattctgctcagc gtctcttaaa gactaaaaga 24360
 cccgcgttt ttccccctcg ggcccaaaaa cccacgtcat cgccagcatg agcaaggaga 24420
 ttcccccccc ttacatgtgg agctatcagc cccagatggg cctggccgcg gggccgccc 24480
 aggactactc cagcaaaatg aactggctca ggcggggccc ccacatgatc tcacgagtt 24540
 acggcatccg agcccaccga aaccagatcc tcttagaaca ggcggcaatc accgcccacac 24600
 cccggcgcca actcaaccccg cccagttggc cccggccca ggtgtatcag gaaactcccc 24660
 gcccgaccac agtccctccctg ccacgcgacgc gggaggccga agtccctcatg actaactctg 24720
 gggtacaatt agcggggcgaaatccaggtacg ccaggtacag aggtcgccgc gtccttact 24780
 ctccccggag tataaagagg gtgatcattc gaggccgagg tatccagctc aacgacgagg 24840
 cggtagctc ctcaaccgtt ctcagacctg acggagtctt ccagctcggaa ggagcgggccc 24900
 gctcttcctt caccactcgc caggcctacc tgaccctgca gagctcttcc tgcagccgc 24960
 gctccggggg aatcggcaact ctccagttcg tggaaagagtt cgtccctcc gtcacttca 25020
 acccgtttc cggctcaccc ggacgcgtacc cggacgcctt cattcccaac tttgacgcag 25080
 tgagtgaatc cgtggacggc tacgactgat gacagatggt gcccgggtga gagctcggt 25140
 ggcacatctc catcaactgcc gccagcctcg ctgtacgct cgggaggcga tcgtgttcag 25200
 ctactttgag ctgcccggacg agcaccctca gggaccggct cacgggttga aactcgagat 25260
 tgagaacgcg cttgagtctc acctcatacgtc cgccttcacc gcccggccctc tccctggtaga 25320
 aaccgaacgc gggatcaacta ccatcacccct gttctgcattc tgccccacgc ccggattaca 25380
 tgaagatctg tgggtgtcatc tttgcgtca gtttaataaa aactgaactt tttgcgtac 25440
 cttcaacgcc acgcgttgtt ttccttgcg aaaaaacccc aggagtccctt aacttacaca 25500
 tagcaaaacc cttgtatccc accatagaaaa aacaacttagc ctttcaatt ggaaaagggt 25560

taacaatttc tgctacagga cagttggaaa gcacagcaag cgtacaggac agcgctacac 25620
 cacccctacg tggtatttcc cctttaaagec tgacagacaa cggttaaca ttaagctatt 25680
 cagatccccct gcgtgtggta ggtgaccaac ttacgtttaa ttttacttct ccactacgtt 25740
 acgaaaatgg cagtcttaca ttcaactaca ctctcccat gacactaata aacaacagtc 25800
 ttgctattaa cgtcaataacc tccaaaggcc tcagtagtga caacggcaca ctcgctgtaa 25860
 atgttactcc agattttaga tttAACAGCT ctggtgccct aacttttgc atacAAAGTC 25920
 tatggacttt tccaacccaaa actcctaact gtaccgtgtt taccgaaagt gactccctgc 25980
 tgagtctttg cttgactaaa tgccggagctc acgtacttgg aagcgtgagt ttaagcggag 26040
 tggcaggaac catgctaaaaa atgacccaca ctctgttac cggtcagtt tcgtttgatg 26100
 acagtggtaa actaatattc tctccacttg cgaacaacac ttgggggtgtt cgacaaagcg 26160
 agagtccgtt gcccaacccca tccttcaacg ctctcacgtt tatgccaac agtaccattt 26220
 attctagagg agcaagtaac gaacctcaaa acaattatta tgtccagacg tatcttagag 26280
 gcaacgtgcg aaagccaatt ctactaactg ttacctacaa ctcagttaat tcaggatatt 26340
 ccttaacttt taaaatggat gctgtcgcca atgaaaaatt tgccactcct acatctcgt 26400
 tttgctatgt tgcagagcaa taaaacccctg ttaccccacc gtctcgcccc tttcagatga 26460
 aacgagcggag agttgatgaa gacttcaacc cagtgatccc ttatgacccccc ccatacgctc 26520
 ccgtcatgcc cttcattact ccgcctttta cctccctcgga tgggttgcag gaaaaaccac 26580
 ttggagtgtt aagtttaaac tacagggate ccattactac acaaaaatggg tctctcacgt 26640
 taaaactagg aaacggcctc actctaaaca accagggaca gttAACATCA actgctggcg 26700
 aagtggagcc tccgctcaact aatgctaaca acaaacttgc actagectat agcgaaccat 26760
 tagcagtaaa aagcaaccgc ctaactctat cacacaccgc tccccttgc atcgctaata 26820
 attctttagc gttgcaagtt tcagagccta tttttgtaaa tgacgatgac aagctagccc 26880
 tgcagacagc cgccccccctt gtaaccaacg ctggcacccct tcgtttacag agcgctgccc 26940
 cttaggatt gggtgaaaat actcttaaac tgcgttttc taaaacccctg tatttgcaaa 27000
 atgattttct tgcatttagcc attgaacgccc ccctggctgt agcagccgca ggtactctga 27060
 ccctacaact tactcttcca taaaagacta acgatgacgg gctaacacta tccacagtcg 27120
 agccatTAAC tggaaaaaac ggaaacccat gcttgcaaat atcgccctt ttagtgttcc 27180
 aaaacaacgg cctttcgctt gctattaccc ccccgctgcg tttgtttaac agcgaccccg 27240
 ttcttggttt gggcttcaact tttcccttag ctgtcacaAAA caacccctc tccctaaaca 27300

tgggagacgg agttaaactt acctataata aactaacgc caatttggtt agggattac 27360
aatttgaaaa cggtgcgatt gccgtaacgc ttactgccga attacctttg caatacacta 27420
acaaacttca actgaatatt ggagctggcc ttcgttacaa tggagccagc agaaaactag 27480
atgtaaacat taaccaaataa aaggcttaa cttgggacaa cgatgcagtt atccccaaac 27540
taggatcggtt cttacaattt gaccctaattt gcaacatcgc tggttatccct gaaaccgtga 27600
agccgcaaac gttatggacg actgcagatc cctcgctaa ctgctcagtg taccaggact 27660
tggatgccag gctgtggctc gctttgtta aaagtggcga catggtgcat ggaagcattg 27720
ccctaaaagc cctaaaaggg acgttgcataa atcctacagc cagctacattt tccattgtga 27780
tatattttta cagcaacggg gtcaggcgta ccaactatcc aacgtttgac aacgaaggca 27840
ccttagctaa cagcgccact tggggatacc gacaggggca atctgctaac actaatgtga 27900
ccaatgccac tgaattttatg cccagctcaa gcaggtaccc cgtgaataaa ggagacaaca 27960
ttcaaaaatca atcttttca tacacctgta ttaaaggaga ttttgctatg cctgtcccgt 28020
tccgtgtAAC atataatcac gcccggaaag ggtattccct taagttcacc tggcgcggt 28080
tagccaatca ggcctttgat attccttgcgtt gttcattttc atacatcaca gaataaaaaa 28140
ccacttttc attttaaattt ctttttattt tacacgaaca gtgagacttc ctccacccctt 28200
ccatttgaca gcatacacca gcctctcccc cttcatagca gtaaaactgtt gtgaatcagt 28260
ccggtatttg ggagttaaaa tccaaacagt ctcttggtg atgaaacgtc gatcagtaat 28320
ggacacaaat ccctgggaca ggtttccaa cggttgggtg aaaaactgca caccgcctta 28380
caaaacaaac aggttcaggg tctccacggg ttatctcccc gatcaaactc agacagggt 28440
aagggtcggtt ggtgttccac taaaccacgc aggtggcgct gtctgaacct ctggcgat 28500
ctcctgtgag gctggtaaga agtttagattt tccagtagcc tcacagcatg tatcatcagt 28560
ctacgagtgc gtctggcgca gcagcgcatc tgaatctcac tgagattccg gcaagaatcg 28620
cacaccatca caatcaggtt gttcatgatc ccatacgatc acacgctcca gccaaagctc 28680
atccgttcca acagcgccac cgctgttccg tccaaacctt ctttaacata aatcaggtgt 28740
ctggcccgta caaacatgct acccacatac agaacttccc ggggcaggcc cctgttccacc 28800
acctgtctgtt accaggaaaaa cctcacattt atcagggage catagatggc cattttaaac 28860
caatttagcta ataccggcccc accagctcta cactgaagag aaccgggaga gttacaatga 28920
cagtgaataa tccatctctc ataacccttg atggtctgat gaaaatctag atctaacgtg 28980

gcacaacaaa tacacacttt catatacatt ttcataacat gttttccca ggccgttaaa 29040
 atacaatccc aatacacggg ccactcctgc agtacaataa agctaataca agatggata 29100
 ctcctcacct cactgacact gtgcatttc atatttcac attctaagta ccgagagttc 29160
 tcctctacag cagcaactgct gcggtcctca caaggtggta gctggtgatg atttagggg 29220
 gccagtctgc agcgataccg tctgtcgct tgcatcgtag accaggaacc gacgcacctc 29280
 ctgcgtacttg tggtagcaga accacgtccg ctgccagcac gtctccacgt aacgcccgtc 29340
 cctgcgtcgc tcacgctccc tcctcaatgc aaagtgcac cactttgtt atccacacag 29400
 atccctctcg gcctccgggg tggatgcacac ctcaaaccctt cagatgtctc ggtacagttc 29460
 caaacacgta gtgaggggcga gttccaacca agacagacag cctgatctat cccgacacac 29520
 tggaggtgga ggaagacacg gaagaggcat gttattccaa gcgattcacc aacgggtcga 29580
 aatgaagatc ccgaagatga caacggtcgc ctccggagcc ctgatggat ttaacagcca 29640
 gatcaaacgt tatgcgatcc tccaagctat cgatcgccgc ttccaaaaga gcctggaccc 29700
 gcacttccac aaacaccagc aaagcaaaag cactattatc aaactcttca atcatcaagc 29760
 tgcaggactg tacaatgcct aagtaatttt cgtttctcca ctcgcaatg atgtcgccgc 29820
 agatagtctg aaggttcatc ccgtgcaggg taaaaagctc cgaaagggcg ccctctacag 29880
 ccatgcgtag acacaccatc atgactgcaa gatatcgccgc tcctgagaca cctgcagcag 29940
 atttaacaga tcaaggtcag gttgctctcc gcgatcaca atctccatcc gcaaggtcat 30000
 ttgcaaaaaaa ttaaataaaat ctatgccac tagatctgtc aactccgcata taggaaccaa 30060
 atcaggtgtg gctacgcagc acaaaagttc cagggatggt gccaaactca ctagaaccgc 30120
 tccccgagtaa caaaactgat gaatgggagt aacacagtgt aaaatgtcata accaaaaatc 30180
 actaaggtgc tcctttaaaa agtccagtagt ttctatattc agtccgtgca agtactgaag 30240
 caactgtgcg ggaatatgca caacaaaaaa aatagggcgg ctcagataca tggtagctt 30300
 aaataaaaaag aatcattaaa ctaaagaagc ttggcgaacg gtgggataaa tgacacgctc 30360
 cagcagcaga caggcaacccg gctgtcccg ggaaccgcgg taattcat ccgaatgatt 30420
 aaaaagaaca acagaaactt cccaccatgt actcggttgg atctcctgag cacacagcaa 30480
 tacccttc acattcatgt cccgcacaga aaaaaacgt cccagatacc cagcggggat 30540
 atccaaacgac agctgcaaag acagcaaaac aatccctctg ggagcgatca caaaatcctc 30600
 cggtaaaaaa agcacatatac tattagaata accctgttgc tggggcaaaa aggcccggcg 30660
 tcccagcaaa tgcacataaa tatgttcatc agccattgcc cctgttacc gcgtaatcag 30720

ccacgaaaaa atcgagctaa aattcaccca acagcctata gctatatata cactccgccc 30780
 aatgacgcta ataccgcacc acccacgacc aaagttcacc cacacccaca aaacccgca 30840
 aaatccagcg ccgtcagcac ttccgcaatt tcagtctcac aacgtcaatt ccgcgcgcct 30900
 tttcacatc ccacacacac ccgcgcctt cgccccgccc tcgcgcacc ccgcgtcacc 30960
 gcacgtcacc ccggcccccgc ctgcgtcctc cccgctcatt atcatattgg cacgtttcca 31020
 gaataaggta tattattgtat gatg 31044

<210> 6
 <211> 34115
 <212> DNA
 <213> simian adenovirus SV-39

<400> 6
 catcatcaat ataacaccgc aagatggcga ccgagttAAC atgcaaATGA ggtgggcgga 60
 gttacgcac ctttgtcttg ggaacgcgga agtggggcgcg gcgggtttcg gggaggagcg 120
 cggggcgggg cgggcgtgtc gcgcggcggt gacgcgcgg ggacccggaa attgagtagt 180
 ttttattcat ttgcagaATT tttctgtaca tttggcgcg aaaactgaaa cgaggaagt 240
 aaaagtgaaa aatgccgagg tagtcacccgg gtggagatct gacctttgcc gtgtggagtt 300
 taccgcgtga cgtgtgggt tcgggtctta tttttcaCT gtggtttcc gggtacggtc 360
 aaaggtcccc attttatgac tccacgtcag ctgatcgta gggtattta tgccctcag 420
 accgtcaaga gccactctt gagtgccggc gagaagagtt ttctcctccg cgttccgcca 480
 actgtgaaaa aatgaggaac ttcttgctat ctccgggct gccagcgcacc gtagccgcgc 540
 agctgttggaa ggacattgtt accggagetc tggagacga tcctcagggtg atttctcaCT 600
 tttgtgaaga ttttagtctt catgatctct atgatattga tccgggtgtt gaggggcaag 660
 aggatgaatg gctggagtct gtggatgggt ttttccgga cgctatgtcg cttagaggctg 720
 atttgcacc acctcacaac tctcacactg agcccgagtc agctgtattt cctgaattgt 780
 catcaggtga acttgacttg gcttgttacg agactatgcc tccggagtcg gatgaggagg 840
 acagcgggat cagcgatccc acggcttta tggctctaa ggctgattgtc atactaaaag 900
 aagatgatga tggcgatgtat ggatttcgac tggacgctcc ggctggtccg gggagagact 960
 gtaagtcctg tgaataccac cgggatcgta ccggagaccc gtctatgtg tggctctgt 1020
 gttatctccg tcttaacgct gctttgtct acagtaagtg ttttgcatt ttttaccctg 1080
 tggcttgggtt gagtttattt tttctgtgt ctcatagggt gttgtttattt ataggctcg 1140

tttcagatgt ggaggaacct gatagtacta ctggaaatga ggaggaaaag ccctccccgc	1200
cgaaactaac tcagcgctgc agacctaata tttgagacc ctcggccca ggtgtgtcat	1260
cccgaaacg tgctgtgtt aattgcata gaaatattt ggaagagccc actgaacctt	1320
tggacttgta cttaaagcga ccccgccccgc agtagggcgc ggtgccagtt tttctctct	1380
agcttccggg tgactcagtg caataaaaat ttcttgca acaggtgtat gtgtttactt	1440
tacggcggg aaggattag gggagtataa agctggaggg gaaaaatctg aggctgtcag	1500
atcgagttag aagttccatg gacttgtacg agagcctaga gaatctaagt tcttgcac	1560
gtttgctgga ggaggcctcc gacagaacct cttacattt gaggttctg ttcggttccc	1620
ctctgagtcg cttttgcac cgggtgaagc gagagcacct gacggaattt gatggcttt	1680
tagagcagct gcctggactg tttgattctt tgaatctcg ccaccggacg ctgcttagagg	1740
agaggctttt tccacaattt gactttctt ctccagggcg tctgtgttca gcgcttgc	1800
ttgctgtaca tctgttgac agatgaaacg agcagacgca gctcagccccg ggttacactc	1860
tggacttctt gacgctatgc ctatggaaatcag tcggaatcag gagggggagg aagctgtacg	1920
ggcgcttggt ggagaggcat ccgtctctgc gccagcagcg tctgcaagct caagtgtgc	1980
tgaggcggga ggtatctggaa gccatttcgg aggaggagag cggcatggaa gagaagaatc	2040
cggagcggg gctggaccct ccggcggagg agtagggggg ataccggacc cttttcctga	2100
gttggctttt ggggcggtgg ggggcgttc tgggtacgt gaggatgaag agggcgcca	2160
acgcggtcag aagagggagc attttgagtc ctgcactttc ttggctgtatg taaccgtggc	2220
cctgatggcg aaaaacaggc tggaggtgg tgggtacccg gaagtatggg aggactttga	2280
gaagggggac ttgcacctgc tggaaaaata taacttttag caggtgaaaa catactggat	2340
gaacccggat gaggactggg aggtggttt gaaccgatac ggcaaggtag ctctgcgtcc	2400
cgactgtcgc taccaggttc ggcacaagg ggtctgcga cgcaacgtgt acctgttggg	2460
caacggcgcc accgtggaga tggtgaccc cagaagggt ggttttggg ccaatatgca	2520
agaaaatgtgc cctgggggtgg tgggtctgtc tgggtgtact tttcatagtg tgaggtttag	2580
cggtagcaat ttgggggtt tggttattac cgcaacact cctgtggtcc tgcataattt	2640
ctactttttt ggcttcagca acacctgtgt ggaaatgagg gtgggaggca aagtgcgcgg	2700
gtgttccccc tacgcttgct ggaaggggggt ggtgagccag ggtaggctt aagtgtctgt	2760
tcacaagtgt atgttggaga gatgcaccc ttggcattcc agtgaggct tccctccacgc	2820

cagcgacaac	gtggcttctg	acaacggctg	cgcctttctt	atcaagggag	ggggtcgcac	2880
ctgtcacaac	atgatatgcg	gccctgggga	tgtcccccca	aagccttacc	agatggttac	2940
ctgcacagat	ggcaagggtgc	gcatgctcaa	gcctgtgcac	attgtgggcc	accggcgcca	3000
ccgctggcca	gagtttgaac	acaatgtgat	gaccggctgt	agcttgtacc	tgggaggcag	3060
gcgaggagtt	ttcttgccca	gacagtgtaa	cctggcccac	tgcaacgtga	tcatggaaca	3120
atccggcgct	acccaggttt	gctttggagg	aatattttagt	ataagcatgg	tggtgtataa	3180
gatccgtgcgc	tacgacgact	gtcggttcgt	tactcgaacc	tgcgactgcg	gagcctctca	3240
cctgtgtaac	ctgactgtga	tggggatggt	gactgaggag	gtgcgactgg	accactgtca	3300
gcactcttgc	ctgcggggagg	agttttcttc	ctcggacgag	gaggactagg	taggtggttg	3360
gggcgtggcc	agcgagaggg	tgggtataa	aggggaggtg	tccgtgacg	ctgtcttctg	3420
tttttcaggt	accatgagcg	gatcaagcag	ccagaccgcg	ctgagcttcg	acggggccgt	3480
gtacagcccc	tttctgacgg	ggcgcttgc	tgcctgggcc	ggagtgcgtc	agaatgttac	3540
cggttcgacc	gtggacggac	gtcccgtgga	tccatctaac	gctgcttcta	tgcgctacgc	3600
tactatcagc	acatctactc	tggacagegc	cgctggcc	gcagccgcca	cctcagccgc	3660
tctctccgcc	gccaagatca	tggctattaa	cccaagcctt	taaaaaaaaa	tatccgtgga	3720
cacccctcagcc	ctggagcttt	accggcgaga	tctagctaa	gtgggtggacc	aactcgcagc	3780
cgtgagccaa	cagttgcagc	tgggtgcac	ccgagtgag	caactttccc	gccccccca	3840
gtaaccgc当地	aaattcaata	aacagaattt	aataaacagc	acttgagaaa	agtttaaact	3900
tgtgggtgac	tttattcctg	gatagctggg	gggagggAAC	ggcgggaacg	gtaagacctg	3960
gtccatcggt	cccggtcggt	gagaacacgg	tggatTTTTT	ccaagacccg	atagaggtgg	4020
gtctgaacgt	tgagatacat	gggcatgagc	ccgtctcggt	ggtggaggtt	ggcccaactgc	4080
agggcctcgt	tttcaggggt	ggtgttgtaa	atgatccagt	cgtaggcccc	ccgctggcg	4140
tgggtctgga	agatgtcctt	cagcagcaag	ctgatggcaa	cgggaagacc	cttgggttag	4200
gtgttgc当地	agcggttgag	ttggggagggg	tgcattgcgg	gactgtgag	gtgcattttg	4260
gcctggatct	tgagggttggc	tatgttgc	cccagatcgc	gcctgggatt	catgttatgc	4320
aagaccacca	gcaccgagta	accggtgca	cggggaaatt	tgtcgtgcag	cttggaaagg	4380
aaagcgtgga	agaatttgg	gaccctcgg	tgcggccct	ggtttccat	gcactcatcc	4440
atgatgtgg	cgatggccc	ccgggaggca	gcctggcaa	aaacgttgcg	ggggccgtg	4500
acatcgtagt	tgtggtcctg	ggtgagttca	tcataggaca	ttttgacaaa	gcgcggcag	4560

agggtccccag	actgggaaat	gatggttcca	tccggtccgg	gggcgttagtt	gcccctcgac	4620
atttgcattt	cccaggcattt	gatttcagag	ggagggatca	tgtcaacctg	ggggggcgtg	4680
aaaaaaaaatgg	tctctggggc	gggggtgatg	agctgggtgg	aaagcaggtt	gcgcaagagc	4740
tgtgacttgc	cgcagccggt	gggcccgtag	atgacagcta	tgacgggttg	cagggtgtag	4800
ttagagac	tacaactgcc	atcatccctc	aaaagcgggg	ccacactgtt	taaaagttct	4860
ctaacatgt	agttttcccg	cactaagtcc	tgcaggagac	gtgaccctcc	tagggagaga	4920
atccaggaa	gcgaagcaaa	gttttaagt	ggcttgaggc	catcgccaa	ggcaagttc	4980
ctgagagttt	gactgagcag	ttccagccgg	tcccagagct	cggtaacgtg	ctctacggca	5040
tctcgatcca	gcagacccctcc	tcgtttccgg	ggttggggcg	gctctggctg	tagggaatga	5100
ggcggtgggc	gtccagctgg	gccatggtgc	ggtcctccca	tgggcgcagg	gttctcttca	5160
gggtggtctc	ggtcacggtg	aatgggtggg	ccccgggctg	ggcgtggcc	agggtgcgc	5220
tgaggctgag	gcccgtggtg	gcgaaccgtt	gttttcgtc	tccctgcaag	tcagccaaat	5280
agcaacggac	catgagctca	tagtccaggc	tctctgcggc	atgtccctttg	gcgcgaagct	5340
tgcctttgga	aacgtccccg	cagttgagc	agagcaagca	tttagcgcg	tagagtttg	5400
gcgc当地	cacggattcc	gggaataag	catccccacc	gcagttggag	caaacggttt	5460
cgcattccac	cagccaggc	agctgaggat	cttttggc	aaaaaccaag	cgcgcgcgt	5520
ttttttgt	gcgc当地	cctcgggtct	ccatgaggcg	gtgc当地	tcgggtacga	5580
agaggctg	gggtctccg	tagacggagg	tcagggcgcg	ctcctccagg	gggtcccgc	5640
ggtcctcggc	gtagagaaac	tcgc当地	ctgacataaa	cgc当地	caggctagga	5700
cgaatgaggc	gatgtggaa	gggtaccgg	cgatcgat	gaggggtcg	gtttttcca	5760
aggtgtgcag	gcacatgtcc	ccctcgccg	cttccaaaaa	tgtgattggc	ttgttaggtgt	5820
aagtcacgt	atcctgtcct	tccgcggggg	tataaaagg	ggcgttccc	ccctcctcgt	5880
cactcttcc	cgttcgtcg	tcgc当地	ccagctgtt	gggtacgtaa	acgc当地	5940
aggcgggcat	gacctgtgcg	ctgaggtgt	cagttctat	atacgaggaa	gatttgcgt	6000
cgagcgc当地	cgtggagatg	cccttgaggt	gtc当地	catttggtca	gaaaacacaa	6060
tctgtcggtt	atcaagctt	gtggcaaaag	acccgtagag	ggcgttggag	agcaacttgg	6120
cgtggagcg	ctgggtttgg	ttttttccc	ggc当地	ttccttggcc	gc当地	6180
gctggacgta	ctccctggcc	acgcacttcc	acccgggaaa	aacggccgtg	cgctcgccg	6240

gcaccagcct cacgctccat ccgcgggtgt gcagggtgat gacgtcgatg ctggggcca 6300
 cctctccgcg caggggctcg ttggtccagc agaggcgacc gcccggcgca gagcagaagg 6360
 ggggcagggg gtcaagcagg cgctcgccg ggggtcgcc gtcgatggta aagatggcgg 6420
 gcagcaggtg tttgtcaaag taatcgatct gatgccccgg gcaacgcagg gcggttccc 6480
 agtccccgcac cgccaaaggcg cgctcgatcg gactgagggg ggcgcggcag ggcattggat 6540
 gcgtcagggc cgaggcgtac atgcccaga tgtcatagac gtaaaggggc tcctccagga 6600
 cgccgaggtg ggtggggtag cagccccccc cgccggatgc ggcccgtacg tagtctaga 6660
 gctcgcgca gggggccaga aggtggcgcc tgaggtgagc gcgctgggc ttttcatctc 6720
 ggaagaggat ctgcctgaag atggcgtggg agtgtggagga gatgggtggc egctaaaaaa 6780
 tggtaagcg ggctcgccggc agacccacgg cctcgccat aaagtggcg taggactctt 6840
 gcagcttttc caccaggag gcggtgacca gcacgtcccg agcgcagtag tccagggttt 6900
 cccgcacgat gtcataatgc ttttcctttt tttccctcca gaggtctcggt ttgaagagat 6960
 actcttcgcg gtctttccag tactcttggg gaggaaaccc gtttctgtt ccacgtaag 7020
 agcccaacat gtaaaaactgg ttgacggcct gatagggaca gcatcccttc tccacggca 7080
 gcgagtaggc cagggcgccg ttgcgcaggg aggtgtgagt cagggcaaag gtgtcgccg 7140
 ccataacttt tacaaaactgg tacttaagt cccggcgatc gcacatgcct cgctcccgat 7200
 ctgagtagtc tgtgcgttt ttgtgcgtgg ggtaggcag ggagtaggtg acgtcgtaa 7260
 agaggatttt gccacatctg ggcataaagt tgcgagagat tctgaagggg ccggcacct 7320
 ccgagcggtt gttgatgact tggcagcca ggagaatttc gtcgaagccg ttgatgttg 7380
 gccccacgac gtagaactct atgaaacgcg gagcgcccg cagcaggggg cactttcaa 7440
 gttgctggaa agtaagtcc cgccgcgtca cgccgtgtt cgtgcggctc cagtccctcca 7500
 ccgggtttcg ctccacaaaa tccgtccaga tgtggtcgac tagcaagagc tgcagtcgt 7560
 cgcgaaattt gcgaaatttt ctgcccgtgg cttgtttctg ggggttcaag caaaaaaagg 7620
 tgtctgcgtg gtgcgcggcag gctcccgac cgagctcgcg agccagattt agggccagca 7680
 gcaccagagc cggctcaccg gtgattttca tgacgaggag aaagggcacc agctgttttc 7740
 cgaacgcgcc catccagggtg taggtctcca cgtcgtaggt gagaaacaga cggtcggtcc 7800
 gcgggtgcga tcccaggggg aaaaacttga tggcgtgcac ccattggag ctctggcgt 7860
 ggatgtgatg gaagtaaaaag tcccgccggc gcgtgaaaca ttcgtgcgtgg ttttgcataa 7920
 agcggccgca gtggcgtcgag cgcgagacgg agtgaaggct gtgaatcagg tgaatcttgc 7980

gtcgctgagg	gggcggcaga	gccaaaaagc	ggagcggaa	cgaccgcgcg	gccacttcgg	8040
cgtccgcagg	caagatggat	gagggttcca	ccgttccccg	cccgcgacc	gaccagactt	8100
ccgcccagctg	cggcttcagt	tcttgcacca	gctctcgag	cgtttcgtcg	ctggcgaat	8160
cgtgaatacg	gaaagtgtcg	ggtagaggcg	ggaggcggtg	gacttccagg	aggtgtgtga	8220
gggcggcag	gagatgcagg	tggtaattga	tttcccacgg	atgacggtcg	cgggcgtcca	8280
aggcgaagag	atgaccgtgg	ggccgcggcg	ccaccagcgt	tccgcggggg	gtctttatcg	8340
cgccgcgggg	cgggctcccg	gccccggcg	cggctggga	cccgcgggc	agtcggcag	8400
cggcacgtcg	cggtggagct	cgggcagggg	ctggtgctgc	gcgcggagct	gactggcaaa	8460
ggctatcacc	cggcgattga	cgtcctggat	ccggcggcgc	tgcgtgaaga	ccaccggacc	8520
cgtggtcttg	aacctgaaag	agagttcgac	agaatcaatc	tcggcatcgt	taaccgcggc	8580
ctggcgcagg	atttcggcca	cgtccccgga	gttgtcttga	tacgcgattt	ctggcatgaa	8640
ctggtcgatt	tccctttct	gcaagtctcc	gtgaccggcg	cgttcgacgg	tggccgcgag	8700
atcggtggag	atgcggccca	tgagctggga	aaaggcattt	atgcgcaccc	cgttccacac	8760
tcggctgtac	accacctctc	cgtgaacgtc	gccccggcgc	atcaccaccc	ggcgagatt	8820
gagttccacg	tggcgccgca	aaaccggata	gtttcgagg	cgtgtataca	gatagttgag	8880
ggtgtggcg	gcgtgctcg	ccacaaaaaa	atacatgtac	cagcggccga	gggtcagtc	8940
gttgcgtcg	cccagcgcct	ccagggcttc	catggcctcg	taaaagtcca	cgccaaagtt	9000
gaaaaattgg	ctgttctgg	ccgagaccgt	gagctttct	tccaagagcc	aatgagatc	9060
cgcacggtg	gccctgactt	cgcgttcgaa	agccccgggt	gcctcctcca	ccttttcctc	9120
ctcgacttct	tcgaccgctt	cgggcacctc	ctttcctcg	accaccaccc	caggcggggc	9180
tcggcggcgc	cgccggccga	cgggcaggcg	gtcgacaaa	cgtcgatca	tttccccct	9240
ccgtcgacgc	atggtctcg	tgacggcgcg	accctgtcg	cgaggacgca	gggtgaaggc	9300
gccgcccgcg	agcggaggta	acagggagat	cggggggcg	tgcgtgggg	gactgacggc	9360
gctaactatg	catctgtatca	atgtttcggt	agtgcacctcg	ggtcggagcg	agtcagcgc	9420
ttgaaaatcc	acgggatcg	aaaaccgttc	caggaacgcg	tctagccaaat	cacagtcgca	9480
aggtaagctg	aggaccgtct	cgggggcttg	tctgttctgt	tttccccgcgg	tggtgctgct	9540
gatgaggtag	ttgaagttagg	cgtcttgag	gcggcggatg	gtggacagga	gaaccacgtc	9600
tttgcggccca	gcttgctgta	tccgcaggcg	gtcggccatg	ccccacactt	ctccttgaca	9660

gcggcggagg tcctttagt attctgcat cagccttcc acggcacct cgttttttc	9720
ttcccgctcg ccggacgaga gccgcgtcag gccgtacccg cgctccccct gtggttggag	9780
cagggccagg tcggccacga cgcgcgtcgc cagcacggcc tgctggatgc gggtaggggt	9840
gtccctgaaag tcgtcgagat ccacaaagcg gtggtaacgcg ccagtgttga tggtaggt	9900
gcagttgttc atgacggacc agttacggt ctgggtgcca tggccacgg tttccaggtt	9960
gcggagacgc gagtaggccc gcgtctcgaa gatgtagtcg ttgcagggtcc gcagcaggtt	10020
ctggtagccc accagcagat gcggccggcg ctggcggtag aggggcacc gctgggtggc	10080
gggggcgttg ggggcgagat cttccaaacat gaggcggtga tagccgtaga tgtagcgcga	10140
catccaagtg atgcccgtgg ccgtgggtct ggcgcggggcg tagtcgcga cgcggttcca	10200
gatgtttcgc agcggctgga agtactcgat ggtggggcga ctctgccccg tgaggcgggc	10260
gcagtcggcg atgctctacg gggaaaaaga agggccagtg aacaaccgcc ttccgtagcc	10320
ggaggagaac gcaaggggggt caaagaccac cgaggctcggtt gttcgaaacc cgggtggcg	10380
cccgaaatacg gagggcggtt ttttgcgtt ttctcagatg catcccggtc tgccggagat	10440
gcgtccgaac gcgggggtccc agtccccggc ggtgcctcg gccgtgacgg cggcttctac	10500
ggccacgtcg cgctccaccc cgcttaccac ggcccaggcg gcggtggctc tgccggcg	10560
aggggaaccc gaagcagagg cggtgttggc cgtggaggag ggccagggggt tggctggct	10620
gggggcctg agtcccggcgc ggcacccggc cgtggctctg aagcgcgacg cggcggaggc	10680
gtacgtgccg cggagcaatc tgtttgcga ccgcagcgcc gaggaggccg aggagatcg	10740
agacttgcgt ttccggcg ggagggagtt gcgtcacggg ctggaccggc agagggttct	10800
gagagaggag gactttgagg cggacgagcg cacgggggtg agtccccgcg gggctcacgt	10860
ggcggccgccc aacctggtga gcgcgtacga gcagacggtc aaggaggaga tgaacttcca	10920
gaagagcttc aatcatcactg tgcgcacgt gattgcgcgc gaagagggtgg ccatcgccct	10980
catgcacatcg tgggattttg tggaggcgta cggtcagaac cccagcagca agccgctgac	11040
ggctcagctg ttccctcatcg tgcaacatag tcgagacaac gaaacgttca gggaggccat	11100
gctgaacatt gcagagcctg aggggcgtg gctttggat ctcattaaca tcttgcagag	11160
tatcgtagtg caggagcgtc cgctgacgcg ggccgacaag gtggctgcca tcaactacag	11220
catgtgtcg ctgggcaaattttacgccccg caagatctac aagtctccgt tcgtccccat	11280
agacaaggag gtgaagatag acagcttttacatgcgcacgc ggcgtcaagg tgctgactct	11340
aagcgacgac ctgggggtgtt accgcaacgacccgcataacac aaggcggtga gcccggcc	11400

ccggcgcgag ctgagcgacc gcgagcttt gcacagcctg catcgccgt tgactggc 11460
 cggcagcgcc gaggcgccg agtactttga cgccggagcg gacttgcgt ggcagccatc 11520
 ccgacgcgcg ctggaggcgg ctggcgtcgg ggagtacggg gtcgaggacg acgatgaagc 11580
 ggacgacgag ttgggcatttgc acttgttagcc gttttcggtt agatatgtcg gcaaacgagc 11640
 cgtctgcggc cgccatggtg acggcgccgg gcgcgcggcc ggaccggcc acgcgcgcgg 11700
 cgctgcagag tcagccttcc ggagtgacgc ccgcggacga ctggtccgag gccatgcgtc 11760
 gcatcctggc gctgacggcg cgcaaccccg aggctttcg gcagcagccg caggcaaacc 11820
 gtttgcggc cattttggaa gcggtggtgc cctccagacc caacccacc cacgaaaagg 11880
 tgctgccat cgtcaacgcgc ctggcgaga ccaaggccat ccgcggccac gaggccgggc 11940
 aggtttacaa cgcgctgcta gaaagggtgg gacgctacaa cagctccaac gtgcagacca 12000
 atctggaccg ctgggtgacg gacgtgaagg aggccgtage ccagcgagag cggttttca 12060
 aggaagccaa tctgggctcg ctggtggccc tcaacgcctt cctgagcacg ctgcggcga 12120
 acgtcccccg cggtcaggag gactacgtga actttctgag cgccctccgc ctgatggtgg 12180
 ccgaggtgcc gcagagcgag gtgtaccagt ctggcccaa ctactacttc cagacctccc 12240
 ggcagggcct gcagacggta aacctgacgc aggctttca gaacctgcag ggctttggg 12300
 gggtgcgcgc tccgctgggc gaccgcagca cgggtccag cctgctgacc cccaatgccc 12360
 ggctgctctt gcttctcatt gctccgttca ccgacagcgg ttccatcagc cgcaactt 12420
 acctgggaca cctgctcacc ctgtaccggg aggccatcg gcaaggcgcgg gtggacgagc 12480
 agacgtacca gaaaaatcacc agcgtgagcc gcgcgtggg gcaggaggac acgggcagct 12540
 tggaggcgcac tctgaacttc ctgctgacca accggcggca ggcctacat cccagtgac 12600
 cgctgaacgc ggaggaggag cgcacccgtc gttcgtgca gcagagcacc gcgtgtact 12660
 tgatgcggga aggccctct cccagcgctt cgctggacat gacggcggcc aacatggagc 12720
 catcgttcta cggcgccaaac cgtcccttcg tcaaccggct aatggactat ttgcattggg 12780
 cggcgccct gaacccggaa tactttacta acgtcataatcgtt gacgaccgt tggctgccac 12840
 ctcccgctt ctacacgggg gagttcgacc tcccgaggc caacgacggt ttcatgtggg 12900
 acgacgtgga cagcgtgttc ctgccccggca agaaggaggc gggtgactct cagagccacc 12960
 ggcgcgaccc cgcagacccgt gggcgaccg ggcgcgtc tccgctgcct cgcctgcga 13020
 ggcgcgaccc cgcagacccgt gggcgaccg ggcgcgtc tccgctgcct cgcctgcga 13080

ggtggAACGA TCCGCTGCTC CGTCGGCCCC GCAACAAAAA CTTCCCCAAC AACGGGATAG 13140
 aggatttttgtt agacaaaatg aaccgttgga agacgtatgc ccaggagcat cggagatggc 13200
 aggccaggca acccatgggc CCTGTTCTGC CGCCCTCTCG GCGCCCGCGC AGGGACGAAG 13260
 acgcccacga ttcagccgat gacagcagcg tggatct gggcgggagc gggAACCCCT 13320
 ttgcccacct gcaacctcgc ggcgtgggtc ggcgggtggcg ctaggaaaaaa aaattattaa 13380
 aagcaactac cagagccatg gtaagaagag caacaaaggt gtgtccctgct ttcttccccgg 13440
 tagcaaaaatg cgtcgggcgg tggcagttcc ctccgcggca atggcgttag gcccgggggg 13500
 ttcttacgaa agcgtgatgg cagcggccac cctgcaagcg ccgttggaga atccttacgt 13560
 gcccggcgca tacctggagc ctacggcgg gagaacacgc attcgttact cggagctgac 13620
 gcccctgtac gacaccaccc gcctgtacct ggtggacaac aagtcaagcag atatcgccac 13680
 cttgaactac cagaacgacc acagcaactt tctcacgtcc gtggcaga acagcgacta 13740
 cacggccggcc gaagcggagca cgcaagaccaat taacttggac gaccgctcgc gctggggcgg 13800
 ggacttggaaa accattctgc acactaacat gccaaacgtg aacgagttca tgtttaccaa 13860
 ctcgttcagg gctaaactta tggtggcgca cgaggccgac aaggacccgg tttatgatgt 13920
 ggtgcagctg acgctggcg aggggaactt ttcagagatt atgaccatag acctgtatgaa 13980
 caacgccatt atcgaccact acctggcggt agccagacag caggggggtga aagaaagcga 14040
 gatcggcgac aagtttgaca cgcgcacatt tcgtctggc tgggacccgg agacggggct 14100
 tgtgatggcg ggggtgtaca cgaacgaagc tttccatccc gacgtggtcc tcttgccggg 14160
 ctgcgggggtg gactttaccc acagccggtt aaacaacctg ctaggcatac gcaagagaat 14220
 gccccttcag gaagggttcc agatcctgtt cggggccgtt gaggccgtt acatccggc 14280
 cctgctggac gtgcggcg acgaggagag catcgccaac gcaagggagg cggcgatcag 14340
 gggcgataat ttgcggcgcc agccccagggc ggctccaacc ataaaaccccg ttttggaaaga 14400
 ctccaaaggcg cggagctaca acgtaatagc caacaccaac aacacggctt acaggagctg 14460
 gtatctggct tataactacg cgcacccggaa gaaggggggtt agggcctggaa ccctgctcac 14520
 cactccggac gtgacgtgcgtt cttcagagca ggtctactgg tcgctgcctg acatgtacgt 14580
 ggaccctgtt acgtttcgctt ccacgcagca agtttgcac tacccagtgg tgggagccgg 14640
 gctttatggcg attcacagca agagtttta caacgcggcag ggcgttactt cacagctcat 14700
 tgcgtcagacc accgcctttaa cgcacgtttt caaccgccttc cccgagaacc aaatcctagt 14760
 gcgacccatcca gcgccccacca tcaccaccgtt cagcgagaac gtgcggcgtt taaccgatca 14820

tgggtggacg	acgggtggga	cgacgacgcc	ccgggtgcgg	caagtccaga	tctccgccac	7140
ggagggggcgc	aggcggttgc	aa	ggagggggacg	cagctgccc	ctgtccaggg	7200
ggcccgcgctg	aggtcggcgg	aa	gaagcgttt	caagttact	ttcagaagac	7260
gtgagccagg	tgcacatgg	tt	acttgattt	caggggggtg	ttggaagagg	7320
gtagaggagg	ccgtgtccgc	gc	gggggcccac	caccgtgcc	cgaggaggtt	7380
cgtcgagggc	gagcgccggg	gg	gttagaggc	ggctctgcgc	cggggggcag	7440
ggcacgtttt	cgtgaggatt	cg	ggcagcgtt	tgatgacgag	cccggagact	7500
gacgacgcgc	ggcggtttag	gt	ccctggatg	tgccgtctct	gcgtgaagac	7560
cgggtccctga	acctgaaaga	ga	gttccaca	aatcaatgt	ctgcacatcg	7620
tgcctgagga	tctcctgtac	gt	cgccccag	ttgtcttgat	aggcgatctc	7680
tgcctccactt	cttcctcgcg	ga	gggtcgccg	tggcccgctc	gtccacacgtt	7740
tcgttggaga	tgcgacgcac	ga	gagttgagag	aaggcggttga	ggccgttctc	7800
cggctgtaca	ccacgtttcc	ga	aggagtcg	cgcgctcgca	tgaccacctg	7860
agttccacgt	ggcgggcgaa	ga	cggtcgtag	tttctgaggc	gctgaaagag	7920
gtgggtggcga	tgtgctcgca	ga	cgaaagaag	tacatgatcc	agcggcgac	7980
ttgatgtctc	cgatggcttc	ga	gacgtctcc	atggccttgt	agaagtcgac	8040
aaaaattggg	agttgcgggc	gg	ccaccgtt	agttcttctt	gcaggaggcg	8100
gcgaccgtgt	cgcgcaccc	ct	gtctcgaaa	gcgccccgag	gacgtctgc	8160
ggctccctct	cttccagggg	ca	cgggttcc	tccggcagct	ctgcacggg	8220
cgacgtcg	gtctgaccgg	ca	ggcggttcc	acgaagcg	cgatcatctc	8280
cgacgcattgg	tctcggtgac	gg	cggttcc	tttcgag	gtcgagttc	8340
ccgcgcagag	cgcgcaccc	ca	gggggtt	aagtggttag	ggccgtcg	8400
gcgcgtacga	tgcatttat	ca	attgtctgc	gtaggcactc	cgtgcaggga	8460
tgcaggtcga	cgggatccga	ga	acttctct	aggaaagcg	ctatccaatc	8520
ggtaagctga	ggacgggtgg	cc	gtctggggg	gcgtccgcgg	gcagttggga	8580
ctgatgtgt	aattaaagta	gg	cggttcttc	aggcggcgga	tggtggcgag	8640
tctttgggcc	cggcctgtt	aa	tgcgacagg	cgctcgcca	tgcccccaggc	8700
cagcgacgca	ggtctttgt	gt	agtcgttgc	atcagtctct	ccacccgaac	8760
ccctgtctg	ccatgcgagt	cg	agccgaac	ccccgcaggg	gctgcagcaa	8820

gccacgaccc	tctcggccag	cacggcctgt	tggatctgcg	tgagggtgg	ctggaaagtgc	8880	
tccaggtcca	cgaagcggtg	ataggcccc	gtgttgcatt	tgttaggtgc	gttggccatg	8940	
acggaccagt	tgacgacttg	catgccgggt	tgggtgcatt	ccgtgtactt	gaggcgcgag	9000	
taggcgcggg	actcgaacac	gtatgcgtt	catgtgcgt	ccagatactg	gtagccaacc	9060	
aggaagtggg	gaggcggttc	tcggcacagg	ggccacgcga	ctgtggcg	ggcgccgggg	9120	
gacaggtcgt	ccagcatgag	gcgtatggtag	tggtagatgt	agcggagag	ccaggtgtatg	9180	
ccggccgagg	tggtcgcggc	cctggtaat	tcgcggacgc	ggttccagat	gttgcgcagg	9240	
gggcgaaagc	gctccatgg	gggcacgc	tgcggcgta	ggcggcgca	atcttgtacg	9300	
ctctagatgg	aaaaaaagaca	ggggcggtcat	cgactccctt	ccgtatgcg	gggggtaaaag	9360	
tcgcaagggt	gcggcgccgg	ggaaccccccgg	ttcgagacccg	gccggatccg	ccgctcccg	9420	
tgcgcctggc	cccgcatcca	cgacgtccgc	gtcgagaccc	agccgcgacg	ctccgc	9480	
atacggaggg	gagtctttt	gtgtttttt	gtatgcac	ccgggtgc	ggcagatgc	9540	
acctcagacg	cccaccacca	ccgcccgcggc	ggcagtaaac	ctgagcggag	gcggtgacag	9600	
ggaggaggag	gagctggctt	tagacctgga	agagggagag	gggctggccc	ggctgggagc	9660	
gccgtccccca	gagagacacc	ctagggttca	gctcgtaggg	gacgccaggc	aggctttgt	9720	
gccgaagcag	aacctgttta	gggaccgcag	cggtcaggag	gcggaggaga	tgcgcatttg	9780	
caggtttcgg	gcgggttagag	.agctgagggc	gggcttcgat	cgggagcggc	tcctgagggc	9840	
ggaggatttc	gagcccgacg	agcgttctgg	ggtgagcccg	gcccgcgc	acgtctcg	9900	
ggccaacctg	gtgagcgcgt	acgacagac	ggtgaacgag	gagcgcact	tccaaaagag	9960	
ctttaacaat	cacgtgagga	ccctgatgc	gagggaggag	gtgaccatcg	ggctgatgc	10020	
tctgtggac	ttcggtggagg	cctacgtgc	gaacccggcc	agcaaacc	tgacggccc	10080	
gctgttcctg	atcgtgc	acagccgc	caacgagacg	ttccgcgacg	ccatgttga	10140	
catcgccggag	cccgagggtc	gctggcttt	ggatctgatt	aacatctgc	agagcatcg	10200	
ggtgcaggag	aggggcctca	gtttagcg	caaggtggcg	gccat	taact attcgatgc	10260	
gagcctgggg	aagttctacg	ctcgcaagat	ctacaagac	ccttacgtgc	ccatagacaa	10320	
ggaggtgaag	atagacagct	tttacatgc	catggcgct	aaggtgc	cgctgagcga	10380	
cgacctcggc	gtgtaccgt	acgacaagat	ccacaaggcg	gtgagcgc	gccgcccggcg	10440	
ggagctgagc	gacagggagc	tgtgcacag	cctgcagagg	gcgc	tggcg	ggcgccgggg	10500

cgaggagcgc gaggcttact tcgacatggg agccgatctg cagtggcgtc ccagcgcg 10560
 cgccttggag gcggcggtc accccgacga ggaggatcg gacgatttg aggaggcagg 10620
 cgagtacgag gacgaagct gaccggcag gtgttgttt agatcgacg gccggcgac 10680
 gggggccaccg cggatcccgc acttttggca tccatgcaga gtcaaccttc gggcgtgacc 10740
 gcctccgatg actgggcggc ggccatggac cgcattatgg cgctgactac cgcacaaaa 10800
 gaggctttta gacagcaacc ccaggccaac cgttttcgg ccatttttggaa agcggtgg 10860
 ccctcccgca ccaacccac acacgagaaa gtcctgacta tcgtgaacgc cctggtagac 10920
 agcaaggcca tccgcccgcga cgaggcggtc ttgatttaca acgctctgtt ggaacgggtg 10980
 gcgcgctaca acagcactaa cgttcagacc aatctggatc gcctcaccac cgacgtgaag 11040
 gagggcgtgg ctcagaagga gcggtttctg agggacagca atctgggttc tctggtagca 11100
 ctcaacgcct tcctgagcac gcagccggcc aacgtgcccc gcgggcagga ggactacgtg 11160
 agcttcatca gcgcctgtgag gctgctggtg tccgaggtgc cccagagcga ggtgtatcag 11220
 tctggggccgg attacttctt ccagacgtcc cgacagggtc tgcaaacggt gaacctgact 11280
 caggccttta aaaacttgca aggcatgtgg ggcgttaagg ccccggtggg cgatcgagcc 11340
 accatctcca gtctgctgac cccaaacact cgcctgtgc tgccttgc tgcgcgcgttc 11400
 accaacagta gcactatcag cctgtactcg tacctgggtc atctcatcac tttgtaccgc 11460
 gagggccatcg gtcaggctca gatcgacgag cacacatatc aggagatcac taacgtgagc 11520
 cgggcctgg gtcaggaaga taccggcgc ctggaagcca cgttgaactt tttgtctaacc 11580
 aaccggaggc aaaaaatacc ctccccatggt acgttaagcg ccgaggagga gaggattctg 11640
 cgatacgtgc agcagtccgt gagtctgtac ttgatgcggg agggccgcac cgcttccacg 11700
 gcttagaca tgacggctcg gaacatggaa ccgtcccttt actccgcac ccggccgttc 11760
 attaacgcgc tcatggacta cttccatcgac gcggccgcac tgaacggggta gtacttcacc 11820
 aatgccatcc tgaatccgca ttggatgcggc ccgtccgggt tctacaccgg cgagtttgac 11880
 ctgcccgaag ccgacgcacgg ctttcttgg gacgacgtgt ccgacagcat tttcacgcgg 11940
 ggcaatcgcc gattccagaa gaaggaggc ggagacgagc tccccctctc cagcgtggag 12000
 gcggccctcta ggggagagag tccctttccc agtctgtttt ccgcacgcag tggtcgggt 12060
 acgcgcggcgc ggttgcggg ggagagcgcac tacctgaacg acccccttgcg cggccggcgt 12120
 aggaagaaaa atttcccaa caacgggggtg gaaagcttgg tggataaaat gaatcggttgg 12180
 aagacctacg cccaggagca gcgaggatgg gaggacagtc agccgcgacc gctgggtccg 12240

ccgcactggc gtcgtcagag agaagacccg gacgactccg cagacgatag tagcgtgttgc 12300
gacctggag ggagcggagc caaccccttt gtcacttgc aacccaaggg gcgttccagt 12360
cgccctact aataaaaaag acgcggaaac ttaccagagc catggccaca gcgtgtgtcc 12420
tttcttctc tctttcttcc tcggcgccgc aqaatgagaa gagcggtgag agtcacgccc 12480
gcggcgtatg agggtccgcc cccttcttac gaaagcgtga tggatcagc gaacgtgccc 12540
gccacgctgg aggcgcctta cgttcctccc agatacctgg gacctacgga gggcagaaac 12600
agcatccgtt actccgagct ggcacccctg tacgatacca ccaagggtga cctgggtggac 12660
aacaagtcgg cggacatcgc ctccctgaat tatcaaaacg atcacagcaa tttctgtact 12720
accgtggtgc agaacaatga ctteaccccg acggaggcgg gcacgcagac cattaacttt 12780
gacgagcgtt cccgctgggg cggtcagctg aaaaccatcc tgcacaccaa catgccccac 12840
atcaacgagt tcatgtccac caacaagttc agggccaggg tgatggttaa aaaggctgaa 12900
aaccagcctc ccgagtacga atggttttag ttcaccattc ccgaggggcaa ctattccgag 12960
accatgacta tcgatctgat gaacaatgcg atcgtggaca attacctgca agtggggagg 13020
cagaacgggg tattggaaaag cgatatcgcc gtaaaatttg ataccagaaa cttccgactg 13080
gggtgggatc ccgtgaccaa gctggtgatg ccaggcgtgt acaccaacga ggctttcac 13140
cccacatcg tgctgctgcc ggggtgcgggt gtggacttca ctcagagccg tttgagtaac 13200
ctgttaggga tcagaaagcg ccgcaccccttc caagagggt ttcagatcat gtatgaggac 13260
ctggaaggag gtaacattcc aggtttgcta gacgtgccgg cgtatgaaga gagtgtaaaa 13320
cagggcggagg cgcaaggacg agagattcga ggcgacacct ttgccacgga acctcacgaa 13380
ctggtaataa aacctctgga acaagacagt aaaaaacggc gttacaacat tatatccggc 13440
actatgaata ccttgtaccc gagctggttt ctggcttaca actacggggta tcccgaaaag 13500
ggagtgagat catggaccat actcaccacc acggacgtga cctggggctc gcagcaagt 13560
tactggtccc tgccggatat gatgcaagac ccggtcaccc tccgcacccctc cacccaaatc 13620
agcaacattcc cgggtgggtggg caccggatcg ctggccgtcc atgccaagag cttctacaac 13680
gaacaggccg tctactcgca actcattcgc cagtcacccg cgcttacccca cgtgttcaat 13740
cgctttcccg agaaccagat tctggtgccgc cctccgcctc ctaccattac caccgtcagt 13800
aaaaacgttc ccgcacccctc acatcacgga accctgccgc tgccgcagcag tatcagtgg 13860
gttcagcgcg tgaccatcac cgacgcccaga cgtcgaacccgt gtcctacgt ttacaaagct 13920

cttggcgtag tggctctaa agtgcgtct agtcgcacct tctaaacatg tccatcctca 13980
 tctctcccga taacaacacc ggctggggac tgggctccgg caagatgtac ggcggagcca 14040
 aaaggcgctc cagtcagcac ccagttcgag ttccccggca cttccgtct ccctggggag 14100
 cttacaagcg aggactctcg ggccgaacgg cggttagacga taccatagat gccgtgattg 14160
 ccgacgcccc cccgtacaac cccggaccgg tcgctagcgc cgcctccacc gtggattccg 14220
 tgatcgacag cgtggtagct ggccgtcggt cctatgtcg ccgcaagagg cggctgcattc 14280
 ggagacgtcg ccccacccgc gccatgctgg cagccaggc cgtgtgagg cggggccgg 14340
 gggtaggcag aagggttatg cgcgcgcgt cgcgcacgc cgcgcgggg agggcccgcc 14400
 gacaggctgc cgcgcaggct gctgcgcaca tcgctagcat ggcgcaccc aggagaggga 14460
 acgtgtactg ggtgcgcgt tctgtgacgg gagtccgagt gccggcgcgc agccgaccc 14520
 cccgaagtta gaagatccaa gctgcgaaga cggcggtact gagtctccct gttgttatca 14580
 gccaacatg agcaagcgca agttaaaga agaactgctg cagacgctgg tgcctgagat 14640
 ctatggccct cccgcgtga agcctgacat taagccccgc gatatcaagc gtgttaaaaa 14700
 gcggggaaaag aaagaggaac tcgcggtggt agacgatggc ggagtggaa ttatttaggag 14760
 tttcgccccg cgacgcagggttcaatggaa agggcggcgg gtacaacgcgt ttttgaggcc 14820
 gggacccgcgt gtagttttt ccccccggaga gccgtcggtc gtttaggggtt tcaaaaggca 14880
 gtacgacgag gtgtacggcg acgaggacat attggAACAG gcggctcaac agatcgaggaa 14940
 atttgcctac ggaaagcggt cgcgtcgca agacctggcc atcgcttttag acagcgccaa 15000
 ccccacgcgc agcctcaaac ctgtgacgt gcagcagggtg ctccccgtga gcgcgcac 15060
 ggacagcaag aggggaataa aaagagaaat ggaagatctg cagccacca tccagctcat 15120
 ggtccctaaa cggcagaggc tggaagaggt cctggagaaa atgaaagtgg acccaagcat 15180
 agagccggac gtcaaagtca ggccgatcaa agaagtggcc cctggctcg gggtgcagac 15240
 ggtggatatc cagatccccg tcacgtcagc ttgcggccgt gttggaaacgc 15300
 aacggaaaacc cctgcccgcgt tcgggtaccag ggaagtggcg ttgcaaaaccg accccctggta 15360
 cgaatacgcc gcccctcgccgtcagaggcgt accccgtcgta cttggcccccgcgt 15420
 catgccagaa tatgcgtgc atccgtctat cctggccacc cccggctacc ggggagtgtac 15480
 gtatcgccccg tcaggaaccc gcccggaaac ccgtcgccgc cccgcgtccct gtcgtgtct 15540
 ggccccccgtg tcgggtgcgcgt gcttaacacg ccggggaaag acagttacca ttcccaaccc 15600
 gcgcgtaccac cctagcatcc ttatgtact ctggccgtttt gcagatggct ctgacttgcc 15660

gcgtgcgcct tcccgttccg cactatcgag gaagatctcg tcgtaggaga ggcattggcgg 15720
 gtagtggtcg ccggcgggct ttgcgcaggc gcatgaaagg cggaatttttta cccgctctga 15780
 taccataat cgcgcgcgc atcggtgcca taccggcgt cgcttcagtgc gccttgcaag 15840
 cagtcgtaa taaataaaacg aaggcttttgc cacttatgtc ctggtcctga ctatttatg 15900
 cagaaaagacg atgaaagaca tcaattttac gtcgtggct ccgcggcacg gctcgccgc 15960
 gctcatgggc accttggaaacg acatcgacac cagtcagctc aacggggcgc ctttcaatttgc 16020
 ggggagcctt tggagcggca ttaaaaaactt tggctccacg attaaatcctt acggcagcaa 16080
 agccttggaaac agtagtgctg gtcagatgct ccgagataaa ctgaaggaca ccaacttcca 16140
 agaaaaagtgc gtcaatgggg tggtagccgg catccacggc gcggtagatc tggccaaacca 16200
 agcggtgcag aaagagatttgc acaggcggtt ggaaagctcg cgggtgccgc cgcagagagg 16260
 ggatgagggtg gaggtcgagg aagttagatg agaggaaaag ctgccccccgc tggagaaaagt 16320
 tcccggtgcg cctccgagac cgcagaagcg acccaggccca gaactagaag aaactctgg 16380
 gacggagagc aaggagccctc cctcgtaacga gcaaggcttgc aaagaggggcgc cctctccacc 16440
 ctacccaatg acaaaaaccga tcgcgcctat ggctcgccgc gtgtacggga aggactacaa 16500
 gcctgtcacg cttagagctcc ccccgccgc accgcccgc cccacgcgc cgcaccgttcc 16560
 ccccccctgt ccggctccgt cggcgccgacc cgtgtccgc cccgtcgccgc tgccctctgcc 16620
 agccgcgcgc ccagtggccgc tggccactgc cagaaacccc agaggccaga gaggagccaa 16680
 ctggcaaagc acgctgaaca gcatcggtgg cctggagtg aaaagcctga aacgcccgc 16740
 ttgttattat taaaagtgttta gctaaaaaaaaat ttccgttgtt atacgcctcc tatgttaccg 16800
 ccagagacgc gtgactgtcg ccgcgcgcgc cgcttcagatggccaccc catcgatgat 16860
 gccgcagtgg tcttacatgc acatcgccgg gcaggacgc tcggagatcc tgagccccgg 16920
 tctcggtgcag ttccgcgcgc ccacccgacac ctacttcagc ttggaaaca agtttagaaa 16980
 cccacccgtg gccccccaccc acgtatgtaac cacggaccgc tcgcaaaggc tgaccctgc 17040
 ttttgtgccc gtagaccggg aggacacccgc gtacttttac aaagtgcgtt acaacgttggc 17100
 cgtaggggac aaccgagtgc tggacatggc cagcacctac tttgacatcc ggggagtgtct 17160
 ggatcgccgtt cccagtttta agccctactc gggtaaccgcg tacaattccc tggctcccaa 17220
 gggcgctccc aaccctgcag aatggacgaa ttcagacagc aaagttaaag tgagggcaca 17280
 ggccgcctttt gtttagctgtt atgggtgttac agcgattaca aaagagggtttaa 17340

agtaacctta acagactccg gatcaacacc acagtatgca gataaaacgt atcagcctga 17400
 gccgcaaatt ggagaactac agtggAACAG cgatgttgg accgatgaca aaatAGCAGG 17460
 aagagtGCTA aagAAAACAA CGCCCATGTT CCCTTGTAC ggctcatATG ccaggcccac 17520
 taatgaaaaa ggaggacagg caacaccgtc cgctAGTCaa gacgtGCAAa atcccGAATT 17580
 acaattttt GCTCTACTA atgtCGCCAA tacACCaaaa gcAGTTCTAT atgcGGAGGA 17640
 cgtgtCAATT gaagcgccAG acactcaCTT ggtgttcaAA ccaACAGTCa ctGAAGGCAT 17700
 tacaAGTTCA gaggctctAC TGACCCAACA agctgtccc AACCGTCCAA ACTACATAGC 17760
 CTTTAGAGAT aattttattG GTCATGTA CTACAATAGC acaggtAAcA TGGGAGTACT 17820
 ggcaggccAG GCTTCTCAGC TAAATGCACT TGTGACCTG caagacAGAA ATACTGAGCT 17880
 GTCCTACCAA CTCATGTTGG acGCCCTCGG agaccGAGT CGGTACTTT CTATGTGGAA 17940
 CCAAGCTGTG GATAGTTACG ATCCTGATGT aagaATCATA gaaaACCATG GCGTAGAAGA 18000
 TGAATTGCT AATTATTGCT TTCTTTGGG aggcatGGCA gtaaccGACA CCTACTCGCC 18060
 TATAAAGGTT AATGGAGGAG GCAATGGATG ggaAGCCAAT AACGGCGTT TCACCGAAAG 18120
 AGGAGTGGAA ATAGGTTCAg ggaACATGTT TGCCATGGAG ATTAACCTGC AAGCCAACCT 18180
 ATGGCGTAGC TTTCTGTACT CCAATATTGG GCTGTACCTG CCAGACTCTC TCAAAATCAC 18240
 TCCTGACAAC ATCACACTCC CAGAGAACAA AAACACCTAT CAGTATATGA ACGGTCGCGT 18300
 GACGCCACCC GGGCTGGTG ACACCTACGT TAACGTGGC GCGCGCTGGT CCCCCGATGT 18360
 CATGGACAGT ATTAACCCCT TTAATCACCA CCGCAACGCC GGACTCCGCT ACCGTTCCAT 18420
 GCTCCTGGGA AACGGACGCT ACGTGCCCTT CCACATCCAG GTGCCCCAGA ATTCTTTGC 18480
 AATTAAAAAC CTGCTGCTGC TCCCCGGTTC CTACACCTAC GAGTGGAACT TCCGCAAGGA 18540
 CGTGAACATG ATCTTGAGA GCTCGCTGGG CAATGACCTG CGAGTGGACG GGGCCAGCAT 18600
 CGCGCTTCGAC AGCATCAACC TGTACGCCA CTTTTCCCC ATGGCCCACA ACACGGCCTC 18660
 CACCCCTGGAA GCCATGCTGC GCAACGACAC CAACGACCAA TCTTCAACG ACTACCTGTG 18720
 CGCGGCCAAC ATGCTGTACG CCATCCCCGC CAACGCCACC AGCGTGCCCA TCTCCATTCC 18780
 CTCTCGCAAC TGGGCAAGCT TCAAGGGCTG GAGTTTCAAC CGCCTAAAAA CCAAGGAGAC 18840
 CCCCTCGCTG GGCTCCGGGT TCGACCCCTA CTTCGTCTAC TCCGGCTCCA TCCCTACCT 18900
 GGACGGCACC TTCTACCTCA ACCATACTTT CAAAAAGGTG TCAATCATGT TCGACTCTC 18960
 CGTCAGCTGG CCCGGCAACG ACCGTCTGCT GACGCCAAC GAGTTGAAAGA TCAAGCGTTC 19020
 GGTGGACGGT GAAGGGTACA ACGTGGCTCA GAGCAACATG ACCAAGGACT GGTTCCTGAT 19080

tcagatgctc agccactaca acatcggtca ccagggttcc tacgtgcggaaaattacaa	19140
ggaccgcattg tactttttct tcagaaactt ccaacccatg agccgcggaaa ttgttagattc	19200
aacggcttac actaattatac aggatgtgaa actgccatac cagcataaca actcagggtt	19260
cgtggctac atgggaccca ccatgcgaga gggcaggcc taccggcca actatcccta	19320
tccctgatt gggccacccg ccgtgcggcag cctcaacgcag aaaaagtcc tctgcgaccg	19380
ggtgatgtgg aggatcccct tctctagcaa ctcatgtct atgggctcccc tcaccgacct	19440
ggggcagaac atgctgtacg ccaactccgc tcacgcctt gatatgaccc ttgaggtgga	19500
tcccatggat gagccccacgc ttcttatgt tctgtttgaa gtcttcgacg tggtgcgcatt	19560
ccaccagccg caccgcggcg tcatcgaggc cgtctacctg cgcacaccc ttctgcggg	19620
taacgcccacc acctaaagaa gccgatggc tccagcgaac aggagctgca ggccattgtt	19680
cgcgacctgg gctgcgggccc ctacttttgg gcacaccccg acaagcgttt tccggcttc	19740
atgtcccccc acaagccggc ctgtgccttc gtaaacacgg ccggacggga gaccgggggg	19800
gtccactggc tcgccttcgc ctggAACCCG cgtaaccgca cctgctaccc ttgcacccct	19860
tttgtttct ccgacgaaag gctgaagcag atctaccagt tcgagtacga ggggctccctc	19920
aagcgcagcg ctctggcctc cacccccgac cactgcgtca ccctggaaaa gtccacccaa	19980
acggtccagg ggcccccttc ggcgcctgc gggctttct gttcatgtt ttgcacccgc	20040
ttcgtgcact ggcctcacac cccatggat cacaacccca ccatggatct gtcacccgga	20100
gtgccccaca gcatgctca cagccccag gtcgcggccca ccctgcggcg taaccaggaa	20160
cacctgtatc gctttctggg gaaacactct gcctatttt gccgcacccg gcagcgcata	20220
gaacggggca cggcccttcga aagcatgagc caaagagtgt aatcaataaa aaacatTTT	20280
atttgacatg atacgcgtt ctggcgaaaa attaaaaatc gaagggttcg agggaggggt	20340
cctcgtgcggc gctggggagg gacacgttgc gatactggaa acgggcgtc caacgaaact	20400
cggggatcac cagccgcggc agggcacgt ctcttaggtt ctgcctccaa aactgcgcga	20460
ccagctgcag ggctcccatg acgtcggcgcc cgatatactt gaagtcgcag tttagggccgg	20520
agctcccccg gctgttgcgg aacacgggggt tggcacactg gaacaccaggc acggccgggt	20580
tgtggatact ggccaggggcc gtcgggttcgg tcacctccga cgcatccaga tcctcgccgt	20640
tgctcagggc aaacggggtc agcttgcaca tctgcggccc aatctgggt actaggtcgc	20700
gctgtttag gcaatgcgcag cgccatggga tcaggatgcg tcgctgcggcg cggtgcata	20760

tagggtaact cgccgccagg aactcctcca tttgacggaa ggccatctgg gctttgccgc 20820
 cctcggtgta gaatagcccg caggacttgc tagagaatac gttatgacccg cagttgacgt 20880
 cctccgcgca gcagcgggccc tcttcgttct tcagctgaac cacgttgccgg ccccaacgg 20940
 tctggaccac cttggctcta gtgggggtgct cttcagcgc ccgtgtccg ttctcgctgg 21000
 ttacatccat ttccaacacg tgctccttgc agaccatctc cactccgtgg aagcaaaaca 21060
 ggacgccttc ctgctgggta ctgcgtatgc cccatacggc gcattccggtg ggctcccagc 21120
 tcttgcgttt tccccccgcg taggcttcca tgtaagccat aaggaatctg cccatcagct 21180
 cggtaaggt cttctggttt gtgaaggta gccggcaggcc gcgggtgcctcc tcgttcaacc 21240
 aagtttgaca gatcttgccgg tacaccgctc cctggtcggg cagaaactta aaagccgctc 21300
 tgctgtcggtt gtctacgtgg aacttctcca ttaacatcat catggtttcc atacccttct 21360
 cccacgcgtgt caccagtggt ttgctgtcgg ggttcttcac caacacggcg gtagagggc 21420
 cctcgccggc cccgacgtcc ttcatggtca ttctttgaaa ctccacggag ccgtccgcgc 21480
 gacgtactct ggcgcaccggg gggtagctga agcccacctc caccacggtg cttcgccct 21540
 cgctgtcgga gacaatctcc gggatggcg gccggcgggg tgctgccttg cgagccct 21600
 tcttgggagg gagctgagggc gcctccgtgc cgcgtcggtt gctcatctcc cgcaagttagg 21660
 gggtaatggta gctgcctgct tggttctgac ggttggccat tgtatcttag gcagaaagac 21720
 atggagctta tgcgcgagga aactttaacc gccccgtccc ccgtcagcga cgaagatgtc 21780
 atcgtcgaac aggacccggg ctacgttacg ccgcccggg atctggaggg gcctgacccgg 21840
 cgcgacgcta gtgagcggca ggaaaatgag aaagaggagg cctgctacct cctggaaggc 21900
 gacgttttgc taaagcattt cgccaggcag agcaccatag ttaaggaggg cttgcaagac 21960
 cgctccggagg tgcccttggta cgtcggcgctt ctcctccagg cctacggcga gaacctttc 22020
 tcgcctcgag tgccctccgaa gagacagccc aacggcacct gcgagccaa cccgcgactc 22080
 aacttctacc ccgtgttgc cgtaccagag gcgtggcca cctatcacat tttttcaaa 22140
 aaccaacgca tccccctatc gtgccccggcc aaccgcaccc cgcccgatag gaatctcagg 22200
 cttaaaaacg gagccaaacat acctgatatac acgtcgctgg aggaagtgcc caagatttc 22260
 gaggggtctgg gtcgagatga gaagcggcg gcaacgcgc tgcaaaaaa acagaaagag 22320
 agtcagaacg tgctgggtggta gctgggggg gacaacgcgc gtctggccgt cctcaaacgc 22380
 tgcatagaag tctcccaactt cgcctacccc gccctcaact tgccacccaa agttatgaaa 22440
 tcggtcatgg atcagctgct catcaagaga gctgagcccc tggatcccgag ccaccccgag 22500

gcggaaaaact	cagaggacgg	aaagccgtc	gtcagcgacg	aggagctcga	gcggggctg	22560
gaaaccaggg	accccaaca	gttgcagag	aggcgcaaga	tgtatgtggc	ggccgtgctg	22620
gtcacccgtgg	agctggaatg	cctgcacacgg	ttttcagcg	acgtggagac	gctacgcaaa	22680
atcgaaaat	ccctgcacta	caccccccgc	cagggctacg	tccgccaggc	ctgcaagatc	22740
tccaaacgtgg	agctcagcaa	cctggctcc	tacatggca	tccctccacga	gaaccggctg	22800
ggcagagcg	tgctgcactg	cacccgtcaa	ggcgaggcgc	ggcgggacta	cgtgcgagac	22860
tgcacatcacc	tccctccac	cctcacctgg	cagaccgcca	tggcgctctg	gcagcagtgc	22920
ttggaaagaga	gaaacacctaa	agagctagac	aaactcctct	gccgccagcg	gcgcgcctg	22980
tggccgggtt	tcagcgagcg	cacggcgcc	agcgctctgg	cgacatcat	cttccggag	23040
cgcctgatga	aaacccgtca	aaacggcctg	ccggatttca	tcaatcaaag	cattttgcaa	23100
aactccgct	ctttgtcct	ggaacgcctc	gggatcttc	ccgccccatgag	ctggcgctta	23160
ccttctgact	ttgtccccct	ctccctaccgc	gagtgcctc	ccccactgtg	gagccactgc	23220
taccccttcc	aactggccaa	cttctggcc	taccactccg	acccatgga	agacgtaaac	23280
ggagagggtt	tactggagtg	ccactgcccgc	tgcacacctgt	gcacccccc	cagatcgctg	23340
gcctgcaaca	ccgagctact	cagcgaaacc	caggctatag	gtacccctcg	gatccagggg	23400
ccccaggcgc	aagagggtgc	ttccggcttg	aagctcactc	cgccgctgtg	gacccctggct	23460
tacttacgca	aattttgttagc	cgaggactac	cacgcccaca	aaattcagtt	ttacgaagac	23520
caatctcgac	caccgaaagc	ccccctcactc	gcctgcgtca	tcacccagag	caagatcctg	23580
gcccaattgc	aatccatcaa	ccaagcgccgc	cgcgatttcc	ttttgaaaaa	gggtcgaaaa	23640
gtgtacctgg	accccccagac	cgccgaggaa	ctcaacccgt	ccacactctc	cgtcgaaagca	23700
gcccccccgaa	gacatgcccgc	ccaaggaaac	cgccaaagcag	ctgatcgctc	ggcagagagc	23760
gaagaagcaa	gagctgcctc	agcagcaggt	ggaggacgag	gaagagatgt	gggacagcca	23820
ggcagaggag	gtgtcagagg	acgaggagga	gatggaaagc	tgggacagcc	tagacgaggaa	23880
ggaggacgag	ctttcagagg	aagaggcgac	cgaagaaaaa	ccacctgcata	ccagcgccgc	23940
ttctctgagc	cgacagccga	agccccggcc	cccgacgccc	ccggccggct	cactcaaagc	24000
cagccgtagg	tgggacgcca	ccgaatctcc	agcggcagcg	gcaacggcag	cggttaaggc	24060
caaacgcgag	cggcgaaaa	attgctccctg	gccccccac	aaaagcagta	ttgtgaactg	24120
cttgcacac	tgccccggaa	acatctccctt	tgcacacgc	tacccctct	tccatcacgg	24180

tgtggccttc cctcgcaacg ttctctatta ttaccgtcat ctctacagcc cctacgaaac 24240
 gctcgagaa aaaagctaag gcctccctcg ccgcgaggaa aaactccgcc gcegctgccg 24300
 ccgccaagga tccacccggcc accgaagagc tgagaaagcg catcttccc actctgtatg 24360
 ctatcttca gcaaagccgc gggcagcacc cttagcgcga actgaaaata aaaaaccgct 24420
 ccttccgctc gctcacccgc agctgtctgt accacaagag agaagaccag ctgcagcgca 24480
 ccctggacga cgccgaagca ctgtttagca aatactgctc agegtctctt aaagactaaa 24540
 agacccgcgc tttttcccccc tcggccgcca aaacccacgt catgccagc atgagcaagg 24600
 agattccac cccctacatg tggagctatc agccccagat gggctggcc gcgggggccc 24660
 cccaggacta ctccagcaag atgaactggc tcagcggccgg ccccccacatg atctcacgag 24720
 ttaacggcat ccgagccac ccaaaccaga ttctctttaga acaggcggca atcaccgcca 24780
 cacccggcg ccaactcaac ccgcctagtt ggcccggccg ccaggtgtat cagaaaaatc 24840
 cccgccccac cacagtccctc ctgcccacgcg acgcggaggc cgaagtcctc atgactaact 24900
 ctgggttaca attagcgggc gggtccaggt acgccaggtt cagaggtcgg gccgctcctt 24960
 actctcccg gagtataaaag agggtgatca ttcgaggccg aggtatccag ctcaacgacg 25020
 agacggtgag ctccctcaacc ggtctcagac ctgacggagt ctccagctc ggaggagcgg 25080
 gccgctcttc ttccaccact cgccaggccct acctgaccct gcagagctct tcctcgccgc 25140
 cgcgctccgg gggaaatcggc actctccagt tcgttggaaa gtttgttccc tccgttact 25200
 tcaacccctt ctccggctcg cctggacgtt accccggacgc ttccattttttt aactttgacg 25260
 cagtgagtga atccgtggac ggctacgact gatgacagat ggtgcggccg tgagagctcg 25320
 gctgcgacat ctgcacatcgtt gccgtcagcc tcgttgcgttac gtcgggagg cgatcgtt 25380
 cagctacttt gagctgccgg acgagcaccc tcagggtccg gctcacgggt tgaaactcga 25440
 gatcgagaac gcgctcgagt ctgcctcat cgacacccctc accggccgac ctctcctgg 25500
 agaaatccaa cgggggatca ctaccatcac cctgttctgc atctgccccca cggccggatt 25560
 acatgaaagat ctgtgttgc atcttgcgc tcagtttaat aaaaactgaa cttttgccg 25620
 cacccctcaac gccatctgtg atttctacaa caaaaagttc ttctggcaaa ggtacacaaa 25680
 ctgtatttta ttctaaatctt acctcatcta tcgtgctgaa ctgcgcctgc actaacgaac 25740
 ttatccagtg gattgcaaac ggtagtgtgt gcaagtactt ttggggaaac gatatagtta 25800
 gttagaaataa cagccttgc gagcactgca actccctccac actaatcctt tatccccat 25860
 ttgttactgg atggtatatg tgctgtggct ccggtttaaa tcctagttgc tttcataagt 25920

ggtttctaca aaaagagacc ctteccaaca attctgttc tttttcgcc ctatcctact 25980
 gctgttctcc ctctggttac tctttcaaac ctctaattgg tatttttagct ttgatactca 26040
 taatctttat taactttata ataattaaca acttacagta aacatgcttg ttctactgct 26100
 cgcacacatct ttcgtctctc ctcacgccag aacaagtatt gttggcgcag gttacaatgc 26160
 aactcttcaa tctgtttaca tgccagattc cgaccagata ccccatatta cgtggtactt 26220
 acaaaccctcc aaacctaatt cttcattttt tgaaggaaac aaactctgcg atgactccga 26280
 caacagaacg cacacatttc cccacccttc actacaattc gaatgcgtaa acaaaagctt 26340
 gaagctttac aacttaaagc cttagattc tggcttgc acatgtgttag ttgaaaaaag 26400
 taatttagaa gtccacagtg attacattga attgacggtt gtggacctgc cacctccaaa 26460
 atgtgaggtt tcctcccttt accttgaagt tcaaggcgtg gatgcctact gcctcataca 26520
 cattaactgc agcaactcta aatatccagc tagaattttac tataatggac aggaaagtaa 26580
 tctttttat tatttaacaa caagcgctgg taacggtaaa cagttacctg actattttac 26640
 tgctgttgtt gaattttcca cctacagaga aacgtatgcc aagcggcctt acaatttctc 26700
 ataccggtt aacgaccttt gcaatgaaat acaagcgctc gaaactggaa ctgattttac 26760
 tccaaattttc attgctgcca ttgttgtaag ctttaattacc attattgtca gcctagcatt 26820
 ttactgcttt tacaagccca aaaaccctaa gttgaaaaa cttaaactaa aacctgtcat 26880
 tcaacaagtg tgattttgtt ttccagcatg gtagctgcatt ttctacttct cctctgtct 26940
 cccatcattt tcgtctcttc aactttcgcc gcagtttccc acctggAACC agagtgccta 27000
 ccgccttttg acgtgtatct gattctacc ttgttgtt gtatatccat ttgcagtata 27060
 gcctgctttt ttataacaat ctttcaagcc gcccactatt ttacgtgcg aattgcttac 27120
 tttagacacc atcctgaata cagaaatcaa aacgttgcct cttactttg ttggcatga 27180
 ttaagttatt gctgatactt aattattttac ccctaataaa ctgttaattgt ccattcacca 27240
 aaccctggtc attctacacc tgttatgata aaatccccga cactcctgtt gttggcttt 27300
 acgcagccac cgccgctttg gtatttatct ctactgcct tggagtaaaa ttgtatttt 27360
 ttttacacac tgggtggcta catcccagag aagatttacc tagatatcct cttgtaaacg 27420
 cttttcaatt acagecctctg cctccctctg atcttcttcc tcgagctccc tctattgtga 27480
 gctactttca actcaccgggt ggagatgact gactctcagg acattaatat tagtgtggaa 27540
 agaatagctg ctcagcgtca gcgagaaaacg cgagtgttgg aatacctgga actacagcaa 27600

cttaaagagt cccactggtg tgagaaagga gtgtgtgcc atgtaagca ggcagccctt 27660
tcctacatg tcagcggtca gggacatgaa ctgtcttaca ctttgcctt gcagaaacaa 27720
accttctgca ccatgatggg ctctacatcc atcacaatca cccaaacaagc cgggcctgta 27780
gagggggcta tcctctgtca ctgtcacgca cctgattgca tgtccaaact aatcaaaaact 27840
ctctgtgttt taggtgatat ttttaagggtg taaatcaata ataaaacttac cttaaatttg 27900
acaacaaatt tctggtgaca tcattcagca gcaccactt accctcttcc cagctctcg 27960
atgggatgcg atagtgggtg gcaaacttcc tccaaaccct aaaagaaata ttggtatcca 28020
cttccttgc ctcacccaca attttcatct tttcatagat gaaaagaacc agagttgatg 28080
aagacttcaa ccccgtctac ccctatgaca ccacaaccac tcctgcagtt ccctttat 28140
ccccccctt tgtaaacagc gatggcttc aggaaaaccc cccaggtgtt ttaagtctgc 28200
gaatagctaa acccctatata ttcgacatgg agagaaaact agcccttca cttgaaagag 28260
ggttgacaat taccggcgcc ggacaattag aaagtacgca gagcgtacaa accaaccac 28320
cggtgataat taccaacaac aacacactga ccctacgtca ttctcccccc ttaaacctaa 28380
ctgacaatag ctttagtgcta ggctactcga gtccctccg cgtcacagac aacaaactta 28440
catttaaactt cacatcacca ctccgttatg aaaatgaaaa ccttactttt aactatacag 28500
agcctcttaa acttataaat aacagcctt ccattgacat caattcctca aaaggcctta 28560
gtagcgtcg aggctcacta gctgtaaacc tgagttcaga cttaaagttt gacagcaacg 28620
gatccatagc ttttggcata caaaccctgt ggaccgctcc gacctcgact ggcaactgca 28680
ccgtctacag cgagggcgat tccctactta gtctctgtt aacccaaatgc ggagctcact 28740
tcttaggaag tgtaagttt accggtttaa caggaaccat aacccaaatg actgatattt 28800
ctgtcaccat tcaatttaca tttgacaaca atggtaagct actaagctct ccacttataa 28860
acaacgcctt tagtattcga cagaatgaca gtacggcctc aaaccctacc tacaacgccc 28920
tggcggttat gcctaacagt accatatatg caagaggggg aggtggtgaa ccacgaaaca 28980
actactacgt ccaaacgtat ctttagggaa atgttcaaaa accaatttattt cttactgtaa 29040
cctacaactc agtgcgccaca ggatattct tatctttaa gtggactgct cttgcacgtg 29100
aaaagttgc aaccccaaca acctcggttt gctacattac agaacaataa aaccgtgtac 29160
cccaccgttt cgtttttttc agatgaaacg ggcgagagtt gatgaagact tcaacccagt 29220
gtacccttat gaccccccac atgctcctgt tatgcccttc attactccac cttttacctc 29280
ctcgatggg ttgcaggaaa aaccacttgg agtgttaagt taaaactaca gagatcccat 29340

tactacgcaa aatgagtctc ttacaattaa actaggaaac ggcctcactc tagacaacca	29400
gggacaacta acatcaaccg ctggcgaagt agaacctcca ctcactaacg ctaacaacaa	29460
acttgcactg gtctatagcg atcccttagc agtaaagcgc aacagcctaa ctttatcgca	29520
caccgcctcc cttgttattg ctgataactc ttttagcattg caagtttcag agcctatttt	29580
tataaatgac aaggacaaac tagccctgca aacagccgcg ccccttgtaa ctaacgctgg	29640
cacccttcgc ttacaaagcg ccgccccctt aggcattgca gaccaaacc taaaactcct	29700
gtttaccaac cctttgtact tgcagaataa ctttctcacg ttagccattt aacgaccct	29760
tgccattacc aatactggaa agctggctct acagctctcc ccacccgtac aaacagcaga	29820
cacaggctt actttgcaaa ccaacgtgcc attaactgta agcaacggga ccctaggctt	29880
agccataaaag cgcccactta ttattcagga caacaacttggact tcagagctcc	29940
cctgcgtctt ttcaacagcg acccagtact agggcttaac ttttacaccc ctcttgcgg	30000
acgcgatgag gcgcctactg ttaacacagg ccgcggcctc acagtggat t acgatggttt	30060
aattttaaat cttggtaagg atcttcgtt tgacaacaac accgtttctg tgcgtcttag	30120
tgcgtcttgc ccttacaat acactgatca gtttcgcctt aacgtggcg ctgggctgcg	30180
ttacaatcca gtgagtaaga aattggacgt gaaccccaat caaaacaagg gtttaacctg	30240
ggaaaatgac tacctcattt gaaagctagg aatggatta gttttgtatg gcatggaaa	30300
catagctgtt ttcctcaag ttacatcgcc tgacacctt a tggaccactg ccgaccatc	30360
ccccaaattgt tccatctaca ctgatttaga tgccaaaatg tggctcttgt tggtaaaaca	30420
aggggggtgtg gttcacggtt ctgttgcctt aaaagcattt a aaggaaccc tattgagttcc	30480
tacggaaagc gccattgtta ttatactaca ttttgcataat tatggagtgc gaattctcaa	30540
ttatcccact ttgggcactc aaggcacgtt gggaaataat gcaacttggg gttataggca	30600
gggagaatct gcagacacta atgtactcaa tgcactagca tttatgcctt gttcaaaaag	30660
gtacccaaaga gggcgtggaa gcaagttca gaatcaaact gtgggctaca cttgtataca	30720
gggtgacttt tctatgcctt caccgtacca aatacagtac aactatggac caactggcta	30780
ctccctttaaa tttatggaa gaaactgtttc aagacaacca tttgacatcc catgctgtt	30840
tttctcttac attacggaaag aataaaacaa cttttcttt ttatttctt tttatttac	30900
acgcacagta aggcttcctc cacccttcca ttcacagca tacaccagcc tctccccctt	30960
catggcagta aactgttgc agtcagtcgg gtatggaa gttaaagatcc aaacaqtc	31020

tttggtgatg aaacatggat ccgtgatgga cacaatccc tggcacagg tctccaacgt 31080
ttcgtaaaa aactgcatac cgccctacaa aacaaacagg ttcaggctc ccacgggtt 31140
tctccccat caaaactcaga cagagtaaag gtgcgatgat gttccactaa accacgcagg 31200
tggcgctgta tgaacctc ggtgcactc ctgtgaggct ggtaagaagt tagattgtcc 31260
agcagcctca cagcatggat catcagtcta cgagtgcgtc tggcgacga ggcacatgt 31320
atctcaactga gattccggca agaatgcac accatcacaa tcaggttgc catgatccca 31380
tagctgaaca cgctccagcc aaagctcatt cgctccaaca gcgccaccgc gtgtccgtcc 31440
aaccttactt taacataaat caggtgtctg ccgcgtacaa acatgctacc cgcatacaga 31500
acctcccccggg gcaaaccctt gttcaccacc tgccctgtacc agggaaacct cacatttac 31560
agggagccat agatagccat tttaaaccctt tagctaaca ccgccccacc agctctacac 31620
tgaagagaac cgggagagg taaatgacag taaataatcc atctctcata acccctaattg 31680
gtctgatgga aatccagatc taacgtggca cagcagatac acactttcat atacattttc 31740
atcacatgtt tttcccgaggc cgtaaaaata caatcccaat acacggggca ctcctgcagt 31800
acaataaagc taatacaaga tggtataactc ctcacccatc taacattgtg catgttcata 31860
ttttcacatt ctaagtaccg agagttctcc tctacaacag cactgcccgc gtcctcacaa 31920
ggtgttagct ggtgacgatt gtaaggagcc agtctgcage gataccgtct gtcgcgttgc 31980
atcgtagacc agggaccgac gcacttcctc gtactttag tagcagaacc acgtccgctg 32040
ccagcacgtc tccaaagtaac gccggccct gctcgctca cgctccctcc tcaacgcaaa 32100
gtgcaaccac tcttgtatcc cacacagatc cctctggcc tccggggcga tgcacacctc 32160
aaacctacag atgtctcggt acagttccaa acacgttagt agggcgagtt ccaaccaaga 32220
cagacagcct gatctatccc gacacactgg aggtggagga agacacggaa gaggcatgtt 32280
attccaagcg attcaccaac gggtcgaaat gaagatccc aagatgacaa cggtcgctc 32340
cggagccctg atggaatttca acagccagat caaacattat gcgattttcc aggctatcaa 32400
tcgcggccctc caaaaagagcc tggacccgca cttccacaaa caccagcaaa gcaaaagcgt 32460
tattatcaaa ctcttcgatc atcaagctgc aggactgtac aatgcccacaa taattttcat 32520
ttctccactc gcaaatgtatg tcgcggcaaa tagtctgaag gttcatgccc tgcatattaa 32580
aaagctccga aaggggcgccc tctatagcca tgcttagaca caccatcatg actgcaagat 32640
atcgggctcc tgagacaccc gcaacgatcc aggtcaggtt gctcccgcc 32700
atcgcaatc tccatccgca aagtcatgg caaataatta aatagatctg cgccgactaa 32760

atctgttaac tccgcgctag gaactaaatc aggtgtggct acgcagcaca aaagtccag 32820
ggatggcgcc aaactcacta gaaccgctcc cgagtagcaa aactgatgaa tgggagtaac 32880
acagtgtaaa atgttcagcc aaaaatcaact aagctgctcc tttaaaaagt ccagtacttc 32940
tatattcagt tcgtgcaagt actgaagcaa ctgtgcggga atatgcacag caaaaaaaaat 33000
aggcggtc agatacatgt tgacctaaaa taaaagaat cattaaacta aagaagctg 33060
gcgaacggtg gnatatatga cacgctccag cagcaggcaa gcaaccggct gtccccggga 33120
accgcggtaa aattcatccg aatgattaaa aagaacaaca gagacttccc accatgtact 33180
cggttggatc tcctgagcac agagcaatac cccctcaca ttcatatccg ctacagaaaa 33240
aaaacgtccc agatacccg cggaatatc caacgacagc tgcaaagaca gcaaaacaat 33300
ccctctggga gcaatcacaa aatcctccgg tgaaaaaagg acatacatat tagaataacc 33360
ctgttgcgtgg ggcaaaaagg cccgtcgcc cagcaatgc acataaatat gttcatcagc 33420
cattgccccg tcttaccgcg taaacagcca cgaaaaatc gagctaaaat ccacccaaca 33480
gcctatagt atatatacac tccacccaaat gacgctaata ccgcaccacc cagcaccaaa 33540
gttcacccac acccacaaaa cccgcgaaaa tccagcgccg tcagcacttc cgcaatttc 33600
gtctcacaac gtcacttccg cgcgcctttt cactttccca cacacgcct tcgcccggcc 33660
gcctctcgcc caccccggt caccggacgt caccgcacgt cacccggcc cgcgcctcgct 33720
cctccccgt cattatcata ttggcacgtt tccagaataa ggtatattat tgatgcagca 33780
aaacaatccc tctggagca atcacaaaat cctccggta aaaaaggcaca tacatattag 33840
aataaccctg ttgctggggc aaaaaggccc gtcgtcccag caaatgcaca taaatatgtt 33900
catcagccat tgccccgtct taccgcgtaa acagccacga aaaaatcgag ctaaaatcc 33960
cccaacagcc tatagctata tataactcc acccaatgac gctaataccg caccacccac 34020
gaccggatc caccacacc cacaatcccg gcggaaatcc agcgccgtca gcaattccgc 34080
aatttcagtc tcacaacgtc acttccgcgc gcctttcac ttccacac acggccctcg 34140
cccgcccccc ctcgcgcac cccgcgtcac cccacgtcac cgacgtcac cccggccccg 34200
cctcgctctt cccgcgtcat tatcatatttgc acgtttcc agaataaggt atattattga 34260
tgca 34264

<210> 5
<211> 31044
<212> DNA

<213> simian adenovirus SV-25

<400> 5		
catcatcaat aatacacctt attctggaaa cgtgccaata tgataatgag cggggaggag	60	
cgagggcgccc ccgggggtgac gtgcggtgac gcgggggtggc gcgagggcgg ggccaagggc	120	
gcgggtgtgt gtgtgggagg cgcttagttt ttacgtatgc ggaaggaggt ttataccgg	180	
aagatggta atttggcggt atacttgtaa gtttgtgt aattggcgcg aaaactgggt	240	
aatgaggaag ttgaggttaa tatgtacttt ttatgactgg gcggaaatttc tgctgatcag	300	
cagtgaacct tggcgctga cggggagggt tcgctacgtg acagtaccac gagaaggcgc	360	
aaaggccc tttattgtac tttcagcgt tttcgctggg tatttaaacg ctgtcagatc	420	
atcaagaggc cactctttag tgctggcgag aagagtttc tcctccgtgc tgccacgatg	480	
aggctggtcc ccgagatgta cggtgtttt agcgacgaga cggtgctaa ctcagatgac	540	
ctgctgaatt cagacgcgcg ggaaatttcc aattcgctg tgcttcgccc gccgtcactt	600	
cacgacctgt ttgtgtttt gctcaacgct tagcaacgtg ttatataggg tcaagaagga	660	
gcaggagacg cagtttgcta ggctgttggc cgatactcct ggagttttt tggctctgga	720	
tctaggccat cactctctt tccaagagaa aattatcaa aacttaactt ttacgtctcc	780	
tggtcgcacg gttgtttccg ctgcctttat tacctatatt ttggatcaat ggagcaacag	840	
cgacagccac ctgtcgtggg agtacatgct ggattacatg tcgatggcgc tggaggggc	900	
catgctgcgg aggagggttt gcatttactt gcgggcccgcg cctccgcggc tggaccgagt	960	
ggaggaggag gacgagccgg gggagaccga gaacctgagg gccgggctgg accctccaac	1020	
ggaggactag gtgtcgagga tgatccgaa gagggacta gtggggctag gaagaagcaa	1080	
aagactgagt ctgaacctcg aaacttttg aatgagttga ctgtgagttt gatgaatcgt	1140	
cagcgtccgg agacaatttt ctggctgaa ttggaggagg aattcaggag gggggactg	1200	
aacctgctat acaagtatgg gtttgaacag taaaaactc actggttggc gccgtggag	1260	
gattttgaaa ccgccttggc cactttgtct aaagtggctc tgccggccggtaaaggtttac	1320	
actatccgcc gcaatgttaa cataaaagaag agtgtttatg ttataggcca tggagctctg	1380	
gtgcaggtgc aaaccgtcga ccgggtggcc tttagttgcg gtatgcaaaa tctggggcccc	1440	
gggggtgatag gcttaaatgg tgtaacattt cacaatgtaa ggtttactgg tgaaagttt	1500	
aacggctctg tgtttgc当地 taacacacag ctgacgctcc acggcgttta cttttttaac	1560	
ttaataaca catgtgtggc gtcgtgggc aggggtgttt tgaggggctg ctgttttac	1620	

ggctgctgga aggccgttgtt gggaaagactt aaaagtgtaa catctgtaaa aaaatgcgtg	1680
tttgagcggt gtgtgttgttgc tttaactgtg gagggctgtg gacgcattag gaataatgcg	1740
gcgtctgaga atggatgttt tcttttgcta aaaggcacgg ctagtattaa gcataacatg	1800
atatgcggca gcggctctgtt cccttcacag ctgttaacctt ggcggatgg aaactgtcag	1860
accttgcgca ccgtgcacat agcgtcccac cagcgcgcg cctggccaac attcgagcac	1920
aatatgctta tgcgttgtgc cgtccacttg ggccttaggc gaggcgtgtt tgcgttgttac	1980
cagtgttaact ttagccatac caagatTTA ctagaacctt ataccccttc tgcgttgtgt	2040
ttaaatgggg tgtttgacat gtcaatggaa ctgtttaaag tgataagata tgatgaatcc	2100
aagtctcggtt gtcgcccattt tgaatgcggaa gctaatttttcc tgaggttgta tcctgttaacc	2160
ctaaacgtta ccgaggagct gaggacggat caccacatgt tgcgttgttgc ggcgcaccgac	2220
tatgaatcca ggcgcgggaa gtgaggtgag gggcggagcc acaaagggtt taaagggcg	2280
tgaggggtgg gtgtgtatgt tcaaaatggaa cgggacgacg gacggcaacg cgtttgagg	2340
tggagtgttc agcccttatac tgacatctcg tttttttttcc tggcaggag tgcgtcagaa	2400
tgttagtgggc tccaccegtgg acggacgacc ggccgcctt gcaattttccg ccacccctcacc	2460
ctatgccacc gtgggatcat cggtggacac tggccggca gctggcgctt ctgcgtccgc	2520
ttctactgtt cggcgttgg cggctgattt tggactgtat aaccaactgg ccactgcagc	2580
tgtggcgtct cgggtctctgg ttcaagaaga tggccctgaat gtgatcctga ctgcgttgttgc	2640
gatcatgtca cgtcgcttgg acgaaactggc tgcgcagata tcccaagctt accccgatac	2700
cacttcagaa tcctaaaata aagacaaaca aatatgttga aaagtaaaat ggcttttattt	2760
gttttttttg gtcggtagg ctcgggttcca cctgtcttgg tgcgttgttgc tttgtgttat	2820
gtttttccaaa acacggtaca gatgggttttgc gatgttcaag tacatggca tgaggccatc	2880
tttgggtgttga agataggacc attgaagagc gtcgttgttgc ggggtgggtgt tgtaaatttac	2940
ccagtcgttag cagggtttctt gggcgtggaa ctggaaagatg tcctttagga gtggctgtat	3000
ggccaaaggcc agggcccttag tgcgttgttgc taaaaggccg ttaagctggg agggatgcgt	3060
gcggggggag atgatatgcg tcttggcttgc gatcttgggg ttagctatgt taccacccat	3120
gtctctgcgg ggggttcatgt tatgaaggac caccacacg gtgtagccgg tgcattttgg	3180
gaacttgcgtca tgcagtttgg agggaaaggc gtggaaagaat ttagagaccc ccttggcc	3240
cccttaggttt tccatgcact catccataat gatggcaatg ggacccctgg cggccgcctt	3300
ggccaaacacg ttttgggggt tggaaaacatc atagtttgc tctagagtgttgc gtcgttgttgc	3360

ggccatctta acaaagcggg gtaggagggt gcccactgg gggatgatag ttccatctgg	3420
gcctggggcg tagttaccct cacagatctg catctcccag gccttaattt ccgagggggg	3480
tatcatgtcc acctgggggg caataaagaa cacggttct ggcgggggat tgatgagctg	3540
ggtgaaaagc aagttacgca gcagttgaga tttgccacag ccggtggggc cgtagatgac	3600
cccgatgacg gggtgcagct ggtagtttag agaggaacag ctgcgcgtcg ggcgcaggag	3660
gggggctacc tcattcatca tgcttctaac atgtttattt tcactcacta agtttgcaa	3720
gagcctctcc ccaccaggataagatcc ttccaggctg ttgaagtgtt tcagcggtt	3780
taggcgtcg gccatggca tctttcgag cgactgacga agcaagtaca gtcggtccc	3840
gagctcggtg acgtgctcta tggaatctcg atccagcaga cttcttggtt gccccgggtt	3900
ggtcgacttt cgctgttaggg caccagccgg tggcgtcca gggccgcgag ggttctgtcc	3960
ttccagggtc tcagcgtccg ggtgagggtg gtctcggtga cggtaaggg atgagccccg	4020
ggctggggcgc ttgcgagggt ggcgttcagg ctcatcctgc tggtgctgaa gggacgtcg	4080
tctccctgtg agtcggccag atagcaacga agcatgaggt cgtagctgag ggactcggcc	4140
gcgtgtccct tggcgcgcag cttcccttg gaaacgtgt gacatttggt gcagtgcaga	4200
cattggaggg cgtagagttt gggggccagg aagaccgact cgggcgagta ggcgtcggt	4260
ccgcactgag cgcagacggc ctcgcactcc actagccacg tgagctcggg tttagcggga	4320
tcaaaaacca agttgcctcc atttttttt atgcgtttct taccttgcgt ttccatgagt	4380
ttgtggcccg cttccgtgac aaaaaggctg tcggtgtc cgtagacaga cttgaggggg	4440
cgtatccca aaggtgttcc gaggtcttcc gcgtacagga actgggacca ctccgagacg	4500
aaggtctgg tccaggctaa cacgaaggag gcaatctgcg aggggtatct gtcgtttca	4560
atgaggggggt ccaccttttc cagggtgtgc agacacaggt cgtcctcctc cgcgtccacg	4620
aaggtgattt gttgttaagt gtaggtcacg tgatctgcac ccccaaagg ggtataaaag	4680
ggggcgtgcc caccctctcc gtcactttct tccgcacgc tgtggaccag agccagctgt	4740
tcgggtgagt aggccctctc aaaagccggc atgatctgg cgctcaagtt gtcagttct	4800
acaaacgagg tggatttgat attcaegtgc cccgcggcga tgctttgat ggtggagggg	4860
tccatctgat cagaaaacac gatcttttg ttgtcaagtt tggtggcgaa agacccttag	4920
agggcgttgg aaagcaactt ggcgatggag cgcagggtct gatTTTCTC ccgatcggcc	4980
ctctccctgg cggcgatgtt gagttgcacg tactccccggg ccgcgcacccg ccactcgggg	5040

cgggacgctg cctttgcaga acagcatccg cgagatccag cgagttacca tcacggacgc 14880
cgtcgctgg acctgtccc acgtctacaa agcattggga atcggtggcc cgcgcttct 14940
gtcgagtcgc actttctaga tgtccatcct catctctccc agcaacaata ccgggttgggg 15000
tctgggcgtg accaaaatgt acggaggcgca caaacgacgg tccccacaac atcccgtgcg 15060
agtgcgcggg cacttttagag ccccatgggg gtcgcacacg cgccggcgca ccggccgaac 15120
caccgtcgac gacgtgatcg atagcgtggt ggccgacgcc cgcaactacc agcccgctcg 15180
atccacggtg gacgaagtca tcgacggcgt ggtggccgac gccaggccct acgcccgcag 15240
aaagtctcggt ctgcgcccgc gccgttcgt aaagcgcccc acggccgcca tgaaagccgc 15300
tcgctctcg ctgcgtcgcg cacgtatcggt gggtegcgcg gccgcacagc ggcgcagccgc 15360
caacgcgcgc gccggccgag tgccgcgcg ggccgcccag caggccgcgc ccgcgcgcgc 15420
cagtcatacc gccccccgac gcgggaatgt gtactgggtc agggactcgg ccaccggcgt 15480
gcgagttccc gtgagaaccc gtcctcctcg tccctgaata aaaagttcta agcccaatcg 15540
gtgttccgtt gtgtgttcag ctcgtcatga ccaaacgcaa gtttaaagag gagctgctgc 15600
aagcgctggt ccccgaaatc tatgcgcccgg cgccggacgt gaaaccgcgt cgctgaaac 15660
gcgtgaagaa gcaggaaaag cttagagacaa aagaggaggc ggtggcgttg ggagacgggg 15720
aggtggagtt tgtgcgtcg ttgcgcgcg gtcggcgagt gaatttggaaagg gggcgcaagg 15780
tgcaacgggt gctgcgtccc ggcacggtgg tgtcttcac cccgggtgaa aaatccgcct 15840
ggaagggcat aaagegcgtg tacgtgagg tgtacggggc cgaagacatt ctggagcagg 15900
cgctggatag aagcggggag tttgcttacg gcaagagggc gaggacgggc gagatcgcca 15960
tccccgttggaa cacttccaac cccacccca gtctgaaacc cgtgacgctg caacaggtgt 16020
tgccgggtgag cgccccctcg cgacgcggca taaaacgcga gggcggcgag ctgcagccca 16080
ccatgcagct cctgggttccc aagaggcaga aactagagga cgtactggac atgataaaaa 16140
tggagcccgaa cgtgcagccc gatattaaaa tccgtcccat caaaagaagtgc ggcgcggaa 16200
tggggcgtgca gaccgtggac atccagattc ccatgaccag cgccgcacag gcggttagagg 16260
ccatgcagac cgacgtgggg atgatgacgg acctgcggcgc agctgctgcc gccgtggcca 16320
gcgcgcgcac gcaaacggaa gcccggatgc agaccgaccc gtggacggag ggcgcggcgtgc 16380
agccggccag aagacgcgtc agacggacgt acggcccccgt ttctggcata atgcccggagt 16440
acgcgcgtgca tccttccatc atccccaccc ccggctaccc ggggcgcacc taccgtccgc 16500

gacgcagcac cactcgccgc cgtcgcccga cggcacgagt cgccaccgcc agagtgagac 16560
gcgttaacgac acgtcgccgc cgccgcttga ccctgcccgt ggtgcgctac catcccagca 16620
ttctttaaaa aaccgcttct acgttgcaga tggcaagct tacttgcga ctccgtatgg 16680
ccgtgcccgg ctacccgagga agatcccccc gacgacggac ttgggagggc agcgggttgc 16740
gcccgcgtcg ggcggttca cggcgctca agggaggcat tctgcccggc ctgatcccc 16800
taatcgccgc agccatcggg gccattcccg gaatcgccag ctagcggtg caggctagcc 16860
agcgccactg attttactaa ccctgtcggt cgccgcgtct ctggcggcag actcaacgcc 16920
cagcatggaa gacatcaatt tctcctctct ggccccggg cacggcacgc ggccgtatat 16980
ggggacgtgg agcgagatcg gcacgaacca gatgaacggg ggccgtttca attggagcgg 17040
tgtgtggagc ggcttgaaaa atttcggttc cactctgaaa acttacggca accgggtgtg 17100
gaactccagc acggggcaga tgctgagggg caagctaaag gacacgcagt ttcaaaaaaa 17160
ggtgggtggac ggcacgtcgg cgggcctcaa cggcgccgtc gacctggcca accaggccat 17220
tcaaaaaggaa attaacagcc gcctggagcc gcccggcag gtggaggaga acctgcccc 17280
tctggaggcg ctgcccccca agggagagaaa gcccggcggg cccgacatgg aggagacgct 17340
atttactaag agcgaggagc cggccatcata cgaggaggcg gtgggttagct cgacgtgcc 17400
gtccctcactg ctgaaggccca ccacctatcc catgaccaag cccatcgctt ccatggcgc 17460
ccccgtggga gtcgacccgc ccatcgacgc ggtggccact ttggacctgc cgcccccga 17520
acccggcaac cgcgtgcctc ccgtccccat cgctccggg gtttctegcc cgcacatccg 17580
ccccgtcgcc gtggccactc cccgctatcc gagccgcaac gccaactggc agaccaccct 17640
caacagtatt gtcggactgg gggtaagtc tctgaagcgc cgctcgctgtt tttaaagcac 17700
aatttattaa acgagtagcc ctgtcttaat ccatcggtgt atgtgtgcct atatcacgcg 17760
ttcagagccct gaccgtccgt caagatggcc actccgtcga tcatgcccga gtggtcgtac 17820
atgcacatcg cccggcagga cgcctcgag tacctgagcc cgggtctggc gcagttgcc 17880
cgtgcgacgg aaacctactt ctcactggc aacaagttca ggaacccac cgtggcgccc 17940
acccacgacg tcaccaccga tgggtcccg cgaactgacaa tccgcttcgt ccccggtggac 18000
aaggaagaca ccgcttactc ctacaaaacc cgcttcacgc tggccgtggg cgacaaccgg 18060
gtgttagaca tggccagtgac ctactttgac atcccgccgc tgatcgacccg cggacactgc 18120
ttcaagcctt actccggcac ggcttacaac tcactggctc ccaaaggggc gcccacaac 18180
agccaatgga acgcccacaga taacgggaac aagccagtgt gttttgtca ggcagctttt 18240

ataggtaaaa gcattacaaa agacggagtg caaatcaga actcagaaaa tcaacaggct 18300
gctgccgaca aaacttacca accagagcct caaattggag tttccacctg ggataccaac 18360
gttaccagta acgctgcgg acgagtgtt aaagccacca ctccccatgtc gccatgttac 18420
ggttcatatg ccaatcccac taatccaaac gggggtcagg caaaaacaga aggagacatt 18480
tcgcttaact ttttcacaac aactgcggca gcagacaata atcccaaagt ggttctttac 18540
agcgaagatg taaaccttca agccccatg actcacttag tatataagcc aacggtggga 18600
gaaaacgtt a cgcgcaga agccctgcta acgcagcagg cgtgtcccaa cagagcaaac 18660
tacataggtt tccgagataa ctttatcggt ttaatgttatt ataacagcac agggaaacatg 18720
ggagttctgg caggtcaggc ctcgcagttt aacgcagttg tagacctgca agatcgaaac 18780
acggaactgt cctatcgat aatgcttagat gctctgggtg acagaactcg atatttctca 18840
atgtgaaatc agggcgttga cagctacat ccagacgtt a gattatcgaaaccatggg 18900
gtggaagacg agctgccaa ttactgtttt ccactccag gcatgggtat ttttaactcc 18960
tacaaggggg taaaaccaca aaatggcggt aatgttaact gggaaagcaaa cggggaccta 19020
tcaaatgcca atgagatcgc ttttagaaac atttttgcca tggaaattaa cctccacgca 19080
aacctgtggc gcagcttctt gtacagcaat gtggcgctgt acctgcccaga cagctataaa 19140
ttcaactcccg ctaacatcac tctgcccccc aaccaaaaaca cctacgagta tatcaacggg 19200
cgcgtcaactt ctccaaaccct ggtggacacc ttgttaaca ttggagcccg atggcgccg 19260
gatcccatgg acaacgtcaa cccctttaac catcaccgga acgcggccct ccgttaccgc 19320
tccatgtgc tggaaatgg acgcgtggtg cttttccaca tacaagtgcc gcaaaaattt 19380
ttcgcgatta agaacctctt gctttgcgc ggctcctaca cttacgagtg gagcttcaga 19440
aaagacgtga acatgattct gcagagcacc ctgggcaatg atcttcgagt ggacggggcc 19500
agcgtccgca ttgacagcgt caacttgtac gccaactttt tcccatggc gcacaacacc 19560
gcttctacctt tggaagccat gctgcgaaac gacaccaacg accagtcgtt taacgactac 19620
ctcagcgcgg ccaacatgct ttatccatt ccggccaacg ccaccaacgt tcccatattcc 19680
atccctccc gcaactgggc ggccttcggg ggatggagct tcacccgcct taaagccaaag 19740
gaaacgcctt cttgggctc cggctttgac ccctactttt tgacttcagg caccattct 19800
tacctggacg gcagctttta cctcaaccac actttcaaac gctctgtccat catgttcgat 19860
tcttcgttaa gttggccggg caacgaccgc ctctgacgc cgaacgagtt cgaaattaag 19920

cgcatgtgg acgggaaagg ctacaacgtg gctcaaagta acatgaccaa agactggttt 19980
 ttaattcaaa tgctcagcca ctacaacatc ggctaccaag gcttctatgt tcccgaggc 20040
 tacaaggatc ggatgtatc ttcttcga aacttcagc ccatgagccg ccaggtgccg 20100
 gatcccacccg ctgcggcta tcaagccgtt cccctgccc gacaacacaa caactcggc 20160
 tttgtgggt acatggccc gaccatgcgc gaaggacage catacccgac caactacccc 20220
 tatccccgtac tcggcgctac cgccgtcccc gccattaccc agaaaaagt tttgtgcgac 20280
 cgcgtcatgt ggcgcatacc ttttccage aactttatgt caatggggc cctgaccgac 20340
 ctcggacaga acatgctta cgctaactcc gccatgccc tggatatgac ttttgaggtg 20400
 gacccatga acgagccac gttgtgtac atgcttttgc aggtgttcga cgtggcaga 20460
 gtgcaccagc cgccacccgg tattatcgag gccgtgtacc tgccgcacccc cttctctgcg 20520
 ggcaatgcca ccacataagc cgctgaacta gctggggggg accccagatc ccatgggctc 20580
 cacggaaagac gaactgcggg ccattgtgcg agacctggc tgccggaccc acttcctggg 20640
 caccttgac aagcggttcc cccgggtcg tgcgcacccaaactcgct gcgcgatcg 20700
 gaataccgccc ggcgagaga cccggggaga gcatggcta gctctgggtt ggaacccccc 20760
 ctgcgtccacg ttttcctgt tgcacccctt tggctttca gaccaacgct tgaagcagat 20820
 ctatgcattt gaatatgagg gtctactcaa gcgaagcgcc ctggcctcccttccgc 20880
 ctgtcttaacc ctggtaaaga gcactcagac gttcaggcc cctcacagcg ccgcctgtgg 20940
 cctttttgt tgcacgcctt tgcacgcctt tgcacgcctt ccggacaccc ccatgaaaa 21000
 caacccacc atggacctcc tgcacgcctt tgcacgcctt tgcacgcctt ccggacaccc 21060
 gcagaccacc ctcctccaaa accagaaaa tctgtacgcc tttctgcaca agcactctcc 21120
 ctactttcgc cgccatcgcc aacaaataga aaatgcaccc gcgtttaca aaactctgt 21180
 acgtttataa aatgaacttt ttattgaact ggaaaacggg tttgtgattt taaaaatca 21240
 aagggggttga gctggacatc catgtggag gccggaaagg tgggtttttt gtaactggc 21300
 ttggggcagcc acttaaaactc tggaaatcaca aacttggca gcggtatttc tggaaatgg 21360
 tcgtgccaca gctggcggtt cagctgaagt gcctgcagaa catcgccggc ggagatctt 21420
 aagtcgcagt ttatctggtt cacggcacgc gcgttgcggt acatgggatt ggcacactga 21480
 aacaccagca ggctgggatt ttgtatgcta gccaggccca cggcgccggt cagtcaccg 21540
 gtgtcttcta tggatggacag cgaaaaaggc gtgactttgc aaagctggcg tcccgccg 21600
 ggcacgcaat ctcctccggta gttgcactca cagcgatgg gcagaagaag atgcttgc 21660

ccgcgggtca tgttagggata ggccgctgcc ataaaagctt cgatctgcct gaaaggctgc 21720
 ttggccttgt gcccttcgggt ataaaaaaaca ccgcaggact tgttggaaaa ggtattactg 21780
 gcgcaagcgg catcgtgaaa gcaagcgcgt gcgtttcgt ttcgttaactg caccacgctg 21840
 cggccccacc ggttctgaat caccttggcc ctgcccgggt tttcctttag agcgcgctgg 21900
 ccggcttcgc tgccccacatc cattccacg acatgctcct tgtaatcat ggcagacccg 21960
 tggaggcagc gcagctcctc gtcatcgctg gtgcagtgtat gctcccacac gacgcagcca 22020
 gtgggctccc acttgggctt ggaggcctcg gcaatgccag aatacaggag aacgtagtgg 22080
 tgcagaaaac gtcccatcat ggtgccaaag gtttctggc tgctgaaggat catcggcag 22140
 tacctccagt cctcgtaag ccaagtgttg cagatcttcc tgaagaccgt gtactgatcg 22200
 ggcataaaagt ggaactcatt gcgctcggtc ttgtcgatct tatacttttc catcagacta 22260
 tgcataatct ccatgcccctt ttcccaggcg caaacaatct tggtgctaca cgggttaggt 22320
 atggccaaag tggttggcct ctgaggcggc gcttgttctt cctttgagc cctctcccga 22380
 ctgacgggggg ttgaaagagg gtgcaccttg gggAACGGCT tgaacacgggt ctggcccgag 22440
 gcgtccccaa gaatctgcat cgggggatttgc ctggccgtca tggcgatgtat ctgaccccg 22500
 ggctccatcca cttegtcctc ctccggactt tcctcggtct ttccggggga cggtaacggga 22560
 gtagggggaa gagcgcggcg cgccttccttc ttggcggca gttccggagc ctgtcttga 22620
 cgactggcca ttgtcttcctc cttaggaaga aaaacaagat ggaagactct ttctcttcct 22680
 cctcgtaac gtcagaaagc gagtcttcca ccttaagcgc cgagaactcc cagcgcatacg 22740
 aatccgatgt gggctacgag actcccccccg cgaacttttc gccgcccccc ataaacacta 22800
 acgggtggac ggactacctg gcccatttggag acgtactgtat gaagcacatc aggccggcaga 22860
 gcgttatcgat gcaagatgtat ctacccggcgc gactcgccgt tccgctggaa gtggcggaaac 22920
 ttagcggccgc ctacgagcga accctcttctt ccccaaagac tcccccaag aggccaggcta 22980
 acggcacctg cgagcctaaac cctcgactca acttcttaccc tgcctttgcc gtgccagagg 23040
 tactggctac gtaccacatt ttttccaaa accacaaaat ccctctctcg tgccgcggcca 23100
 accgcaccaa agccgatcgc gtgtcgac tggaggaagg ggctcgatata cctgagattg 23160
 cgtgtctgga ggaagtccca aaaatctttg aaggctctggg ccgcgacgaa aagcgagcag 23220
 caaacgctct ggaagagaac gcagagagtc acaacagcgc cttggtagaa ctcgaggcgg 23280
 acaacgccag actggccgtc ctcaaaacgggt ccatagaagt cacgcacttc gcctaccccg 23340

ccgttaacct ccctccaaaa gttatgacag cggcatgga ctgcgtgctc ataaagcg 23400
ctcagccctt agacccagag cacgaaaaca acagtgcga aggaaaaccg gtggttctg 23460
atgaggagtt gagcaagtgg ctgtcccca acgaccccgc cacgttggag gaacgaagaa 23520
aaaccatgat ggccgtggtg ctatgttaccg tgcaattaga atgtctgcag aggttcttt 23580
cccacccaga gaccctgaga aaagtggagg aaacgctgca ctacacattt aggcacggct 23640
acgtgaagca agcctgcaag atttccaacg tagaacttag caacctcatc tcctacctgg 23700
ggatcttgca cgaaaaccgc ctggacaaa acgtgcgtc cagcacactg aaaggagaag 23760
ccccccgaga ctatgtgcga gactgcgtgt tcctagcgtc agtgtacacc tggcagagcg 23820
aatggggagt ctggcagcag tgccctggagg acgaaaacct caaagagctt gaaaagctgc 23880
tggtgcgtc cagaagggca ctgtggacca gtttgacga gcgcaccgc gcgcgagacc 23940
tagctgatat tattttcct cccaagctgg tgcaagactt ccgggaagga ctgccagatt 24000
ttatgagtca aagcatcttcaaaaacttcc gtcctttcat cttggAACgc tcggaaatct 24060
tgcccgccac tagctgcgc ctacccacag attttgtgcc tctccactac cgcaatgcc 24120
caccggcgt gtggccgtac acttacttgc ttAAacttgc caactttcta atgttccact 24180
ctgacctggc agaagacgtt agcggcgagg ggctgctaga atgccactgc cgctgcaacc 24240
tgtgcacccc ccaccgcgtct ctagtatgca acactcccct gctcaatgag acccagatca 24300
tcggtacctt taaaatccag ggaccctccg acgcggaaaa cggcaagcag gggctggc 24360
taaaactcac agccggactg tggacctccg cctacttgcg caaatttgc ccagaagact 24420
atcacggcca ccaaattaaa ttttacaaaa accaatcaaa accacccaaa agcgagttaa 24480
cggttcgt cattacgcag agcagcatag ttggcagtt gcaagccatt aacaaagcgc 24540
ggcaagagtt tctcctaaaa aaaggaaaag gggctactt ggaccccccag accggcgagg 24600
aactcaacgg accctccca gtcgcagggtt gtgtccccca tgccgccccaa aaagaacacc 24660
tcgcagtggc acatgccaga gacggaggaa gaggagtggc gcagtgtgag caacagcgaa 24720
acggaggaaag agccgtggcc cgaggggtgc aacggggaaag aggacacggc gggacggcga 24780
agtcttcgccc gaagaactct cgccgcgtcc cccgaagtcc cagccggccg cctccggccca 24840
agatcccgca cacacccgtt gatggatag caagacaaaa aagccgggtt agagaaacgc 24900
tcgccccccgc cagggctacc gtcgcgtggag aaagcacaac aactgcattt ttcgtgcctt 24960
gctccagtgc ggcggagacg tttcggtcac ccgtagatac ttgcctttt acaaagggtt 25020
ggccgtcccccc cgttaacgtcc tccactacta ccgtcactct tacagctccg aagccggacgg 25080

ctaagaaaaac gcagcagttg ccggcgggag gactgcgtct cagcgccga gaacccccag 25140
ccaccaggga gctccgaaac cgcatatttc ccaccctcta cgctatctt cagcaaagcc 25200
ggggcagca gcaagaactg aaaataaaaa accgcacgct gaggtcgctt acccgaagct 25260
gccttatca caagagcgaa gagcagctgc agcgaacctt ggaggacgca gaagcgctgt 25320
tccagaagta ctgcgcgacc accctaaata actaaaaaag cccgcgcgca ggacttcaaa 25380
ccgtctgacg tcaccagccg cgcccaaaa tgagcaaaga gattcccacg ctttacatgt 25440
ggagttacca gccgcagatg ggattagccg ccggcgcgc ccaggattac tccacgaaaa 25500
tgaactggct cagcgcggg ccccacatga tttcccgctt aaacgacatt cgcccccacc 25560
gcaatcagct attgttagaa caggctgtc tgaccgccac gccccgtaat aacctgaacc 25620
ctccccagctg gccagctgcc ctggtgtacc agggaaacgcc tccacccacc agcgtacttt 25680
tgccccgtga cgcccaggcg gaagtccaga tgactaacgc gggcgcgcaaa tttagcggcg 25740
gatcccggtt tcggtacaga gttcacggcg ccgcaccccta tagcccaggt ataaagaggc 25800
tgatcattcg aggcagaggt gtccagctca acgacgagac agtgagctt tcgcttggtc 25860
tacgaccaga cggagtgttc cagctcgccg gctcgcccg ctcttcgttc acgcctcgcc 25920
aggcataacct gactctgcag agctctgcct ctcaagctcg ctccggagga atcggacccc 25980
ttcagttgtt ggaggagttt gtgcctcg tctactttca gcctttctcc ggatcgcgg 26040
gccagttaccc ggacgagttc atccccaaact tcgacgcggt gagtgactct gtggacggtt 26100
atgactgtatg tcgagccgc ttcaagtgtca gtggacaacta cgcggctcaa tcacctggtt 26160
cgttggccgc gccgctgctg cgtggctgc gacttgagct tagctctcaa gtttgtaaaa 26220
aaccctgtccg aaaccgggag cgctgtgcac gggttggagc tagtgggtcc tgagaaggcc 26280
accatccacg ttctcagaaa ctttgtggaa aaacccattt tggtaaaacg agatcagggg 26340
ccttttgtaa tcagcttact ctgcacctgt aaccatgttgc accttcacga ctatttatg 26400
gatcatttgtt ggcgtgaatt caataagtaa agcgaattct taccaagatt atgatgtcca 26460
tgactgttcc tcgcccactat acgatgttgtt gccagtaaac tctcttgctcg acatctatct 26520
gaactgttcc ttttggtccg cacagcttac ttggtactac ggtgacacccg tcctttctgg 26580
ctcaactgggc agctcacacg gaataacact tcacctttt tcgcccgttcc gatacgaaaa 26640
ctacagctgt cgtgcggta cctgcctcca cgtttcaat ctgcagccct gtccacccgac 26700
caaacttgta tttgtcgact ctaagcactt acagctcaac tgcagcattc taggccccag 26760

tatcttgtgg acataacaata aaatcagggtt ggtggaaattt gtctactacc cacccagcgc 26820
ccgcgggtttt ggggaaattc ctttccagat ctactacaac tatcttgcca cacattatgc 26880
aagtcaacag caactaaact tgcaagcacc cttcacgcca ggagagtaact cctgtcacgt 26940
aggctcctgc acagaaaactt ttattcttta caacagatct tctgccattg aacgcttcac 27000
tactaactac tttagaaacc aagttgtgtt tttcaactgac gaaacccta acgtcacccct 27060
ggactgtgca tggttttctc atgacaccgt aacttggact cttacaataa ctctctggct 27120
cgcggtcgat aaccaaagct tgattgttaa aaattttgat ttaacctta ctaaaccctc 27180
tcctcgcga atagttatct ttgctccccc taatccaaaa actacccctt cctgtcaggt 27240
tttgttaag ctttgccaaa caaactttaa gtttgttat ttgcctccgc aatctgtcaa 27300
actcatagaa aaatacaaca aagcggccgt ctggctccct aaaacccctt accactggct 27360
aacctacacg gggctgtttt cactaattgt tttttccctt attaacattt ttatatgttt 27420
cttgccctcc tccttctttt cgcaacacc gttggcccgag aaagacccctt ccttattact 27480
gtagcgcttg ctataaaaaa ccaagagtgg tcaaccgtgc tctcaatcta tttcaattt 27540
ttcattttgt ccttaataact ttctcttatt gtcgttaaca atgatctgga gcattggct 27600
cgccctttttt tggctgctta gtgcaaaagc cactatttt cacaggtatg tggaagaagg 27660
aactagcacc ctctttacga tacctgaaac aattaaggcg gctgatgaag tttcttggta 27720
caaaggctcg ctctcagacg gcaaccactc attctcagga cagaccctt gcatccaaga 27780
aacttatttt aaatcagaac tacaatacag ctgcataaaaa aacttttcc atctctacaa 27840
catctcaaaa ccctatgagg gtatttacaa tgccaagggtt tcagacaact ccagcacacg 27900
gaacttttac ttaatctga cagttattaa agcaatttcc attcctatct gtgagtttag 27960
ctccccagttt ctttctgaaa cctactgttt aattactata aactgcacta aaaatgcct 28020
tcacaccacc ataatctaca atcacacaca atcacccctgg gttttaaacc taaaattttc 28080
tccacacatg ctttcgcaat ttctcacgca agttaccgtc tctaaacataa gcaagcagtt 28140
tggcttttac tttcttcc acgaactgtg cgaaataatt gaagccgaat atgaaccaga 28200
ctactttact tacattgcca ttgggtgtat cgttgggttgc ctttgcattt gtttgggg 28260
gtgtgtttat ttgtacattc agagaaaaat attgcctcg ctgtgcctt gcggttacaa 28320
agcagaagaa agaattaaaaa tctctacact ttattatgt tttccagaaa tggccaaaact 28380
aacgctccta cttttgccttc tcacgcccgt gacgctttt accatcactt tttctgccc 28440
cgccacacactc gaacctcaat gtttgcacc ggttgaagtc tactttgtct acgtgttgc 28500

gtgcgtcggtt agcggttgca gtataacatg ttttaccttt gtttttcgtt agtgcattga 28560
 ctacttctgg gtcagactct actaccgcag acacgcgcct cagtatcaa atcaacaat 28620
 tgccagacta ctgggtctgc catgattgtc ttgtatTTTA ccctgatTTT tttcacctt 28680
 acttgcgtttt gtgatTTTca cttcactcaa ttttggaaaa cgcaatgttt .cgacccgcgc 28740
 ctctccaaacg actggatgtat ggcttttgcattt attgccacgc ttggggcggtt tggacttttt 28800
 agtgggttttgcattt caaatTTAAG actccatggaa cacatggctt tctttcagat 28860
 tttccagttt cacctactcc gcccctcccc ccggccatcg acgtgcctca ggttccctca 28920
 ctttctccat ctgtctgcag ctactttcat ctgtaatggc cgacctagaa tttgacggag 28980
 tgcaatctga gcaaagggtt atacacttcc aacgcgcgtt ggaccgcgaa cgcaaaaaca 29040
 gagagctgca aaccatacaa aacacccacc aatgtaaacg cgggatattt tgtattgtaa 29100
 aacaagctaa gctccactac gagcttctat ctggcaacga ccacgagctc caatacgtgg 29160
 tcgatcagca gcgtaaaacc tttgtattttt taattggagt ttccccattt aaagttactc 29220
 aaaccaaggg tgaaaccaag ggaaccataaa ggtgctcatg tcacctgtca gaatgcctt 29280
 acactcttagt taaaacccata tttgtgttac atgattctat cccctttat taaataaaact 29340
 tactttaaat ctgcaatcac ttcttcgtcc ttgtttttgtt cgccatccag cagcaccacc 29400
 ttccccctttt cccaaacttcc atagcatatt ttccgaaaag aggcttactt tcgcccacacc 29460
 ttaaaggaa cgtttacttc gcttcaagc tctccacga ttttcatTCG agatatgaaa 29520
 cgcgccaaag tggagaagg atttaacccc gtttatccctt atggatattt tactccgact 29580
 gacgtggctc ctccccctt agcctctgac ggttcaagaaaacccacc tgggtcttg 29640
 tccctaaaaaa tatccaaacc tttaactttt aatgcctcca aggctctaag cctggctatt 29700
 ggtccaggat taaaattca agatggtaaa ctgtggggg agggacaagc aattcttgc 29760
 aacctggccgc ttcaaatcac caacaacaca atttcactac gtttggaa cacacttgcc 29820
 ttgaatgaca ataatgaact ccaaaccaca ctaaaatctt catgccttcc taaaatcaca 29880
 gaccagactc tttcccttaa cataggggac agccttgcaa ttaaagatga caaacttagaa 29940
 agcgcttcc aagcgaccct cccactctcc attagcaaca acaccatcg cctcaacgtg 30000
 ggcacccggac tcaccataaa tggaaacgtt ttacaagctg ttcccttaaa. tgctctaagt 30060
 cccctaacta ttccaacaa taacatcagc ctgcgtatg gcagttccctt gacgggtctt 30120
 aacaatgaac tgcaaagcaa cctcacagtt cactccctt taaaactcaa ctccaacaac 30180

tcaatttctc tcaacactct atctccgttt agaatcgaga atggtttcct cacgctctat 30240
ttgggaacaa aatctggctt gctagttcaa aacagtggct taaaagttca agcgggctac 30300
ggcctgcaag taacagacac caatgctctc acattaagat atctcgctcc actgaccatt 30360
ccagactcgg gctcagaaca aggcatctt aaagtaaaca ctggacaggg cctaagtgtg 30420
aaccaagctg gagcgttga aacatcccta ggaggtggat taaaatatgc tgataacaaa 30480
ataaccttg atacaggaaa cggaactgaca ttatctgaaa ataaacttgc agtagctgca 30540
ggtagtggtc taacttttag agatggtgcc ttggtagcca cgggaaccgc attacgcaa 30600
acactgtgga ctacggctga tccgtctccc aactgcacaa ttatacagga ccgcgacaca 30660
aaatttactt tggcgcttac cattagtggg agccaagtgc tggggacggg ttccattatt 30720
ggagtaaaag gcccccttc aagtagcata ccgtcagcta ccgttacagt acaacttaac 30780
tttatttcca acggagccct attgagctcc ttttactta aaggttactg ggggtatcgc 30840
caaggtccct caattgaccc ttacccata attaatgcct taaactttat gccaaactca 30900
ctggcttatac ccccgggaca agaaatccaa gaaaaatgta acatgtacgt ttctactttt 30960
ttacgaggaa atccacaaaag accaatagtt ttaaacatca cttaataaa tcaaaccagc 31020
gggtttcca ttagatttac atggacaaat ttaaccacag gagaagcatt tgcaatgccc 31080
ccatgcactt tttcttacat tgctgaacaa caataaaacta tgtaaccctc accgttaacc 31140
cgccctccgcc cttccatttt atttataaa ccacccgatc cacctttca gcagtaaaca 31200
attgcatgtc agtaggggca gtaaaacttt tgggagttaa aatccacaca ggttcttcac 31260
aagctaagcg aaaatcagtt acacttataa aaccatcgct aacatcgac aaagacaagc 31320
atgagtccaa agcttccggc tctggatcag atttttgttc attaacagcg ggagaaacag 31380
cttctggagg attttccatc tccatctct tcatacgatc caccatgtcc accgtggtca 31440
tctgggacga gaacgacagt tgcatacac ctcataagtc accgggtcgat gacgaacgta 31500
cagatctcga agaatgtcct gtcgcgcct ttcggcagca ctggggccgaa ggcgaaagcg 31560
cccatgttta acaatggcca gcaccccccgtt ctcatacgagg cgcctagttc ttttagcgca 31620
acagcgcatg cgcagctcgta agactggc gcaagaaaca cagcacagaa ccaccagatt 31680
gttcatgatc ccataagcg tgcacacca gccataacta acaaattgtt tcactattct 31740
agcatgaatg tcataatctga tggcaagta aattaaatgg gccccctta tgtaaacact 31800
tcccacgtac aacacctcct ttggcatctg ataattaacc acctcccgt accaaataca 31860
tctctgatta atagtcgccc cgtacactac ccgattaaac caagttgcca acataatccc 31920

ccctgccata cactgcaaag aacctggacg gctacaatga cagtgcaaag tccacaccc 31980
 gttgccatgg ataactgagg aacgccttaa gtcaatagtg gcacaactaa tacaaacatg 32040
 taaaatagtgt ttcaacaagt gccactcgta tgaggtgagt atcatgtccc aggaaacggg 32100
 ccactccata aacactgcaa aaccaacaca tcctaccatc ccccgcacgg cactcacatc 32160
 gtgeatggtg ttcatatcac agtccggaag ctgaggacaa ggaaaaagtct cgggagcatt 32220
 ttcatagggc ggttagtgggt actccttgc ggggttcagt cggcaccggc atctcctcac 32280
 cttctggcc ataacacaca agttgagatc tgatttcaag gtactttctg aatgaaaacc 32340
 aagtgcattc ccaacaatgt atccgatgtc ttccggcccc gcgtcggtag cgcccttgc 32400
 agtacacacg gaacaaccac tcacgcaggc ccagaagac a gtttcccg gacggtgaca 32460
 agttaatccc cctcagtctc agagccaata tagtttctc cacagtagca taggccaaac 32520
 ccaaccagga aacacaagct ggcacgtccc gttcaacggg aggacaagga agcagaggca 32580
 gagggcatagg caaagcaaca gaattttat tccaaactggc cacgtagcac ttcaaacacc 32640
 aggtcacgta aatggcagcg atcttgggtt tcctgatgga acataacacg aagatcaaac 32700
 atgagacgat tctcaagggt attaaccaca gctggattt aatccctccac ggcacattt 32760
 agaaaacacca gcaatacaaa agcccggtt tctccggat ctatcatagc agcacagtca 32820
 tcaattagtc ccaagtaatt ttcccggtt caatctgtt taatttgcag aataatgcc 32880
 tgtaaatcca agccggccat ggcggaaagc tcagataatg cacttccac gtgcattcgt 32940
 aaacacaccc tcatcttgc aatccaaaaa gtcttcttct tgagaaacct gtagtaaatt 33000
 aagaatcgcc aggttaggt cgatgcctac atcccgagc ttcaattctca gcatgcactg 33060
 caaatgatcc agcagatcag aacagaatt agcagccagc tcatccccgg tttccagttc 33120
 cggagttccc acggcaatta tcactcgaaa cgtgggacaa atcggaaataa catgagctcc 33180
 cacgtgagca aaagccgtag ggccagtgca ataatcacag aaccagcgga aaaaagattg 33240
 cagctcatgt ttcaaaaagc tctgcagatc aaaattcagc tcatttcataat aacacagtaa 33300
 agtttgcggg atagtaaccg aaaaccacac gggtcgacgt tcaaaacatct cggcttacct 33360
 aaaaaagaag cacatttta aaccacagtc gcttcctgaa caggagggaaa tatggtgccg 33420
 cgtaaaacca gacgcgccac cggatctccg gcagagccct gataatacag ccagctgtgg 33480
 tttaaacagca aaacctttaa ttccggcaacg gttgaggtct ccacataatc agcgcccaca 33540
 aaaatcccat tcgaaacttg tcgcgttagg gagctaaaat ggccagtgata gccccatggc 33600

acccgaacgc taatctgcaa gtatatgaga gccacccat tcggcggat cacaatca 33660
gtcgagaaa acaacgtata cacccggac tgcaaaagct gttcaggca acgcccgtc 33720
ggtcccttc ggtacaccag caaagcctcg ggtaaagcag ccatgccaag cgcttaccgt 33780
gccaagagcg actcagacga aaaagtgtac tgaggcgctc agagcagcgg ctatatactc 33840
tacctgtgac gtcaagaacc gaaagtcaaa agttcacccg gcgcgcgcga aaaaacccgc 33900
gaaaatccac ccaaaaagcc cgcgaaaaac acttccgtat aaaattccg ggttaccggc 33960
gcgtcaccgc cgcgacac gcccccccg cccccgcgtc ctccccgaaa cccgcgcgc 34020
ccacttccgc gttcccaaga caaaggcgc gtaactccgc ccacctcatt tgcattttaa 34080
ctcggtcgcc atttgcgtt gttatattga tgatg 34115

<210> 7
<211> 44
<212> DNA
<213> Artificial

<220>
<223> primer PSL

<400> 7
gcgacacgcgt ctctatcgat gaattccatt ggtgatggac atgc

44

<210> 8
<211> 36
<212> DNA
<213> Artificial

<220>
<223> primer P5ITR

<400> 8
gcgatattaa atcatcatca ataataacc tcaaac

36

<210> 9
<211> 31
<212> DNA
<213> Artificial

<220>
<223> primer PSXTOP

<400> 9
gataacctagg aacgaggagg atttgatatt g

31

<210> 10
<211> 20

<212> DNA
<213> Artificial

<220>
<223> primer P5XBOT

<400> 10
atgtacgcct ccgcgctcac

20

<210> 11
<211> 31
<212> DNA
<213> Artificial

<220>
<223> primer P5E4

<400> 11
gatcgaattc ccactctgta ccccatctct g

31

<210> 12
<211> 31967
<212> DNA
<213> Simian adenovirus

<220>
<221> CDS
<222> (13796)..(15322)

<220>
<221> CDS
<222> (18257)..(21010)

<220>
<221> CDS
<222> (27192)..(29015)

<400> 12
catcatcata atatacccta tttgggaacg gtgccaatat gataatgagg aggcggggtt 60
aggggtggag tgaggggtggg gtgcggatga cgcgccgcg gggcggggtg ggagtctgac 120
gtggggcgcg gggtggagcg cgaggggtgag ggcggggcga gggcggcggg cgccggcggaa 180
ttgacgtaca cggtagtaag tttgagcggaa aattaagtga attgggcgtg tttttgtaa 240
ctttttgacg tacacggtag taagtttgag cggaaaattaa gtgaattggg cgtgttttt 300
gtaactttt gacgtacacg gtagtaagtt tgagcggaaa tttaagtgaat tgggcgtgtt 360
ttttgttaact ttttgacgta cacggtagta agtttgagcg gaaattaagt gaattggcgcg 420
tgtttttgtt aactttttga cgtacacggt agtaagttt agcggaaaatt aagtgaattt 480

ggcgtgttt ttgttaactt ttgacgtaca cggtagtaag tttgagcgg aattaagtga	540
atgggcgtg tttttgtaa cttttgacg tacacggtag taagttttag cgaaaattaa	600
gtgaattggg cgtgtttttt gtaactttttt ggteattttgcgcgaaaac tgagtaatga	660
ggaagtgaga cggactctgc cctttttac ggttggagg gaaaactgct gatcagcgct	720
gaactttggg ctctgacgca gttttccc tacgtggcag tgccacgaga aggctcaaag	780
tcctcgttt attgtgtgct cagcctttt gagggtattt aaacaccgtc agaccgtcaa	840
gaggccactc ttgagtgca gcgagtagag tttctccctc cgctcgctgcc gcgctgctc	900
agtcttaccg ccaggatgcg aatgcgtccg gagatctca cccggctctg ggaagatgtt	960
ttccagggac ttttagaaatc tgaagacaac tttccccaaac ctccctgagcc ggaggagcta	1020
cctgaggttt cgcttcacga tctgtttgac gtggaggtgg agagccccga cggagatccg	1080
aacgaggaag ctgttgcattt tatgttcccc gactggatga tatctcagag cgagagtgc	1140
gaaggcagtg cggactcggg cgtttctggg gttggaaacc tgggtggaggt ggatctggac	1200
ttgaagtgtt acgaggaagg tttctccct agcgactcag agactgtatga agcctcagaa	1260
gcggaaaggta aagaggagtc tttgtgtggt tatgtgaaga ttaatgggg ggagaacctg	1320
ctgggtttgg actgtccgga ccaacctgga catggctgtc gagcctgtga ctccctccgg	1380
gggaccagcg gaaacccgga agctatctgt gctttgtgtc acatgcgtct gaacgagcac	1440
tgcataataca gtgagtgta ttcatgggtt atttatgggg aaagttgggg gaaagtcttgc	1500
agaaggggaa aagtttaaca tgtcatttttt gtacttgata ggtccagttt cggacgtgca	1560
gggggattct gagtccccctg ctggcccttc ccagccctca ccctgtctt tgaccgcccac	1620
ccccgcaccc gacctagttt gaccaacgcc ctgcccggatg tcctgttagac gacgtgcagc	1680
tgttaattgc atagaagatt tattggcccc tggatgacgag aacgcaccc ttgaccccttag	1740
cctgaaacgc cctaagacat cttgagtgat tatgtgttta ataaaagtgt tgacccttag	1800
atccctgtgtt tattccctgg gcgtgtgcgc gggtatataa agcagctgcg ggctggagtg	1860
ttagtttattt ctgatggagt actggagtgat gctgcagaat taccagagcc tccggcgcct	1920
gctggagtttgcgc gctctgtccca gaacatccac ctgctggagg ttctgttttgcgc gctgcactct	1980
cagtaacgtg gtgtatcgaa tgaagcaaga gtacagctcg cgctttctg agctgttggc	2040
ccgcgtacccg gctgttttttgcgc gctctgtccca tctaggccat cacgtttattt tccaagaagc	2100
tgttagtcaga tattttggattt tttctactcc cgggcgtgcg gtttctgtca ttgccttcat	2160
ctgctttgtg cttagatcgat ggagcgcggca aacccgcctg agcccggggtt acaccctggaa	2220

ctacacctgacc	atgtccctgt	ggagggccat	gctgcggaag	aggagggtct	caggcttc	2280
gccggcgccgg	cctccgcacg	gactggatcc	ggtgctggag	gagtccggagc	tggaggagga	2340
ggagaacccg	agggccggcc	tggaccctcc	ggcggaaatag	tgacggaaacc	ggaggatccc	2400
caagagggtta	ctagtcaggg	gggaggggggg	ccgaagagaa	agcgggatga	agaggaggcg	2460
atggaccccg	acaggtttct	aaaagaactg	actttaagct	taatgtctaa	gagaagaccc	2520
gagacggtgt	ggtggctctga	tttggagaag	gagttccacc	agggggagat	aatctgttg	2580
tacaagtatg	ggttttagca	ggtgaagact	cactggctgg	aagcctggga	ggactggag	2640
atggctttta	acatgtttgc	caagggtggcg	ctgcgccccgg	acactattta	caccgtgact	2700
aagacggtgg	aaatccgcaa	gcctgtgtat	gtgattggca	acggggccgt	ggttcgggttc	2760
cagaccacccg	accgggtggc	ctttaattgc	tgtatgcaga	acctggggccc	gggggtgatt	2820
aatcttaatg	gagtgacctt	ttgcaatgtc	agattcgcgg	gggatggatt	caacgggacg	2880
gtgtttccggc	ccaccaccca	gataacccta	cacgggggtgt	tcttccagca	tgtaggcggg	2940
gcttgtgttag	atacctgggc	gagggcctct	gtgaggggct	gcacctttgt	ggctgttgg	3000
aaagcggtgg	tgggtcgacc	caagagtgtg	ctgtctgtga	agaaatgtgt	gtttgagaga	3060
tgtctgatgg	ccatgggtgt	ggagggccag	ggttaggatcc	gccataacgc	gggctccgag	3120
aataacctgtt	ttgcccctgct	gaagggtacg	gcgaccgtga	agcataaacat	gatctgcggg	3180
gtgggtcaact	cgcagctgct	gacctgtgcg	gatggcaact	gccaggccct	gcgcacggtg	3240
catgtggtgt	cccacccggcg	ccgccccctgg	ccgggttttg	aacataaacat	gctgatgcgc	3300
tgtaccatgc	acctgggcta	ccgcccggc	gtgtttgtgc	cccatcagtg	taacctgacc	3360
cacaccaagg	tgttgctgga	gacggatgt	ttttcgcgag	tgaatctgaa	tgggggttcc	3420
gatctgacta	tggagatgta	caagatagtg	agattttagt	aatcaaagac	ccgttgcgc	3480
ccctgcaagt	gcgggtccaa	tcacacctgagg	atgtatcccg	tgaccctgaa	cgtgacggag	3540
gagctgcgcc	cgaccaccca	gatgctgtcc	tgtctgcga	ccgattacga	aagcagcgt	3600
gaggattaag	aggtgaggggg	cgggggttgc	atgggggtata	aaggtgggggg	aggaggtggg	3660
gagggggaaa	acccaaaatg	agcggatcga	tggaaaggag	cgctgtgagt	tttgaggggcg	3720
gggtgttccag	cccatatctg	acaacccgtc	tcccccctg	ggcaggagtg	cgtcagaatg	3780
tggtgccctc	caacgtggac	ggacgtccgg	tggccccctgc	caactccgccc	actctcacct	3840
acgccacccgt	cggatcgtcg	ctggacacccg	ccgctgcccgc	cgccgcttca	gcccccgtt	3900

ctactgctcg	cggtatggca	gctgatttcg	gactgtatca	gcaactggct	gcgcctcgct	3960
cgtcgctgag	agaagatgat	gccctgtccg	tggtgctgac	ccgcctggag	gagctgtccc	4020
agcagctgca	agagctgtct	gccaaagtgg	atgcacagaa	cgtccccgt	acccaatgaa	4080
taaataaaacg	agacaccgag	tgtgtttgga	aatcaaaatg	tgtttttatt	tgtttttct	4140
ggcgcggtag	gcccttgacc	acctgtcgcg	gtcgtaagg	accttgtgga	tgttttccag	4200
cacccggtag	aggtgggctt	ggatgtttag	gtacatgggc	atgagcccgt	ctcgggggtg	4260
gaggttagcac	cactggaggg	cgtcgctc	gggggtgggt	ttgtagataa	tccagtcgt	4320
gcagggtttt	tgggcattgga	agcggaaagat	gtctttgaga	agcaggctga	tggccagggg	4380
gagggccttg	gtgttaggtgt	tcacaaagcg	gttgagctgg	gagggatgca	tgcgaaaaaa	4440
gatgagatgc	atcttggcct	aatcttgag	gttggcgatg	ttgcggccca	gatcccggcg	4500
ggggctcatg	tttgtcagga	ccaccaggac	ggtgtagccg	gtgcacttgg	ggaatttgtc	4560
atgcaacttg	gaagggaaagg	cgtggaaagaa	cttggagacc	cccttgtggc	cgccgaggtt	4620
ctccatgcat	tcgtccatga	tgtggcgat	gggacccctg	gcggccggccc	tggcgaagac	4680
gttgcgggg	tgggagacgt	cgtagttctg	ttccagggtg	agctcgctgt	aggccatttt	4740
gacgaagcgg	gggagcaggg	tgcccgactg	ggggacgatg	gtaccttcgg	gaccgggggc	4800
gtagttgccc	tgcagattt	gcatctccca	ggccttgate	tccgaggggg	ggatcatgtc	4860
cacctggggc	gcgatgaaga	agacggcttc	cggggcgggg	ttgtatgact	gggaggagag	4920
gaggttgcgg	agcagctgcg	acttggcgc	cccggtggc	ccgttagatga	ccccgatgac	4980
gggttgcage	tggtagttt	aggagctgca	gtgcccgtcc	tgcgcagga	acggggcgac	5040
ctcgatc	atgcttcgt	'cgtgatggtt	ttccctgacg	aggcttgca	agagccgctc	5100
gccgcccagg	gagagaagct	cttccaggct	gcgaaatgc	ttgaggggtt	tgaggccgtc	5160
ggccatggtc	atctttcca	gggactggcg	gagcaggtac	aggcggctcc	agagctcggt	5220
gacgtttct	acggcatctc	gatccagcag	acttcttgg	tgcgggggtt	ggggcggctt	5280
tggctgtägg	ggaccagccg	gtgcgcgtcc	aggaggcga	gggtgacgtc	tttccagggc	5340
cgcagcgttc	gctgtgagggt	ggtctcggt	acggtaagg	gatgcgcgtcc	cggttggcg	5400
ctggccaggg	tcctcttgag	actcatcctg	ctgggtgtgga	agcgggcgtc	ttctccctgg	5460
gagtccggcca	ggtagcattt	gagcatgagg	tctgtatgat	gggcctcgcc	cgctggccc	5520
ttggcgcgca	gcttgccttt	ggagacgtgt	ccgcaggcgg	gacagtgcag	gcacttgagg	5580
gcgttagagct	tggggggccag	gaagacggac	tcgggggagt	aggcgtcgcc	gcccactga	5640

gcccacgtgg tctcgactc gacgagccag gtgagctcg ggtgttgggg atcaaaaacc	5700
agctggcccc cgtgtttttt gatgcgttc ttacctcggt tctccatgag gcggcgtccg	5760
gcttcggtga cgaagaggct gtcgggtcg ccgtagacgg atttgagcgc gcgtgtcc	5820
aggggaatcc cgcgatcctc cgctgcagg aactcgacc actctgagac gaaggcccg	5880
gtccacgcga ggacaaaggga ggcatctgg gacgggtage ggtcggtctc caccaggga	5940
tccacccctt ccagggtgtg caggcagagg tcgttcttcc cccgtccat gaaggtgatt	6000
ggctttaag tgtatgtcac gtgaccgtcg ggtcgcccg tgggttata aaagggggcg	6060
tgcggccct ccccgtaact ttcttccgca tcgtgtggc cgagatccag ctgtcgggt	6120
gagtaggcgc gctggaaaggc gggcatgacc tcggcgctga gggtgtcagt ttccacgaac	6180
gaggtggatt tgatattgac ctgtccggcg gcatgtctt tgacggtggc ggggtccatc	6240
tggtcagaaa agacgatctt ttgttgcac agcttgggtgg cgaacgaccc gtagagggcg	6300
ttggagagca gcttggcgat ggagcgcagg gtctggttt tctcgccgtc ggccgcgtcc	6360
ttggccggcga tggtgagctg gacgtactcg cggccacgc agcgcattc gggaaagacg	6420
gtggccgcgt cgtccggcag gaggcgcacg cgcgcgcgc gtttgtcgag ggtgtatgagg	6480
tccacgtgg tggccacctc gccgcgcagt ggctcggtgg tccagcagag gcgcggccc	6540
ttgcgcgagc agaaggggggg caggacgtcg agctggctt ccgcgggggg gtcggcgctcg	6600
atggtaaga tgcccggtag caggtggcg gtcgaagtatg cgtggcgac cgcggggtcg	6660
gcgagggcgc gttcccagtc cctgaccgcg aggccgcgt cgtaggggtt gagggccgc	6720
ccccagggca tggatgggt gagggccgag gctacatgc cgcacatgtc gtagacgttag	6780
aggggctcgc ggagcgcgc gaggtaggtg gatagcgc gtcgcgcgc gatgtggcg	6840
cgcacgtatg cgtacatctc gtgcgaggg gcgaggaggc cgcctccgag gtcgcgcgc	6900
tgcggcttga cggcccggtt ggtgacctgg cggaaagatgg cgtgcgagtt ggaggagatg	6960
gtggccgcgt gaaagatgtt gaagctggcc tcggggagtc cgcggcggtc gtggacgaac	7020
tgggcgttagg agtcgcgcag cttctgcacg agcgcggcg gtcgcgcgc gtagacgttag	7080
cagtagtcga gggctcgcg gacgaggctcg taacggggctt ttccagatgt	7140
tgcgggttga ggaggtactc ctcgcgtatcc ttccagatctt cttccggccgg aaagccgcgt	7200
tgcgcgcgttca ggtaagaacc cgcacatgttag aagcggttga cggctcggtt gggacagcag	7260
cccttctcga cggccaggga gtggccgtc gcggcccttcc tgagcgcgtt gtgggtgagg	7320

gccaagggtgt	cgcgcacccat	gaccttggagg	aacttggaaacc	tgaagtccgt	gtcgatcgac	7380
gcgcggccgt	cccagagccc	gtatcggtg	cgtttctggc	tgcgggggtt	ggcgaggggcg	7440
aagggtacgt	cggttgaagag	gatcttgcgg	gcgcggggca	tgaagttgcg	ggtgatcctg	7500
aaggggcccg	gcacgtccga	gcgggtgtta	atgacctggg	ccgcgaggac	gatctcgatcg	7560
aagccgttga	tgttgtggcc	gacgatgttag	agctcgacga	agcgccggcg	ccccctgcagc	7620
ttggggggctt	tctttagctc	ctcgtaggtg	aggcagtcgg	gcgagtagag	gcccatctcc	7680
tgtcgccccc	attcgccac	ctgggggttg	gcttgcaaga	agcccccca	gagctgcagg	7740
gcgagctggg	tctggaggcg	gtcgccgttag	tgcggaaact	ttttgcac	cgccatcttc	7800
tcgggggtga	ccacgtagaa	ggtgcggccg	tcctggcccc	aggcgtcccc	gttctcgatcg	7860
cggcgcagac	ggcaggccctc	ctcgacgagg	gcctccccc	cgagagatg	catgactagc	7920
atgaagggg	cgagttgttt	gcgcaggca	cccatccacg	tgttaggtctc	tacgtcgtag	7980
gtgacgaaga	gacgttcgg	gcgaggatgc	gagccgagag	gaaagaagtt	gatctccctgc	8040
caccagccgg	aggagtgggc	gttgcgtgg	tggaaagtaga	agtcacgccc	gcggaccgtg	8100
cattcggtct	gatatttgt	aaagcggcg	cagtactcg	agcgctgcac	gctctgcact	8160
tcctgaacga	gatgcacccg	gcgcggcg	accaggaggc	ggagggggca	gtccagtgg	8220
gcttcggcg	gctgtccctc	agcctcgatc	tgccttctg	cacctgcacg	ctccctgtgt	8280
gggtggagga	cggagggagt	gacgacgccc	cgcgagccgc	aggccatgt	gtcgacgcgc	8340
ggcgccctga	ggctcagcgc	cagggtgcgg	atctgagcgg	cgtccaggga	gtcgaggaag	8400
gcctcgctga	ggtcgacggg	cagcgtccgc	cgggtggactt	gcaggagacg	ggtaaggggcc	8460
ggcgccaggg	gctgtatgtt	cttgcgttc	agcggttcgt	tgggtggaggt	gtcgatggcg	8520
tagagcaggg	cctgaccgcg	ggcggcgacg	atgggtccgc	ggtgcggcg	gttaggtggcg	8580
tattcggggg	ggctcggttac	atcacccgc	tgggcctggc	gcccggcg	agcggggggtt	8640
ctggccccgc	cggcatgggc	ggcagcggca	cgtcgccgcg	gggcctccggc	agcggctgtgt	8700
gtcgagctcg	cagctgactg	gcgtgcgcga	cgacgcggcg	gttgcgttc	tggatgtgcc	8760
tcccgctcg	gaagaccacc	ggtccccgg	ctcgaaacct	gaaagagatg	tcgcacagaat	8820
caatctcgcc	atcggttgc	gcccgcgtac	gcaggatctc	ctgcacgtcg	cccgagttgt	8880
cctggtaggc	gatctcgac	atgaactgg	cgatctcttc	ctccctggagt	tgcggcg	8940
cggcgccgttc	gacgggtggcc	gcgaggatcg	tggagatgcg	agccatgagc	tgggagaagg	9000
cgttgcggcc	gttctcgatc	cacacgcac	tgtagacgac	gttgcggacg	gcgtccccgg	9060

cgcgcatgac cacctgcgcg acgtttagct ccacgtgtcg cgcgaagacg gcgttagttgc 9120
 gcaggcgctg gaagaggtag ttgagggtgg tggcgatgtc ctcgcagacg aagaagtaca 9180
 tgaccaggcg ggcgcagcg tc atctcggtga tgtctccgag ggcttccaag cgctccatgg 9240
 cctcgtagaa gtcgacggcg aagttgaaga actgggagtt ggcgcggcg accgtcagct 9300
 cgtcttgcaaa gagccggatc agctgggcca cggctcccg cacctgcgt tcgaaggccc 9360
 cggcgcttc ttccctctt ggttccctcg cggcctcttc ttccatgacg gtttcccttt 9420
 cctccggttc ctccggcact ggcctccggc ggcgcacggcg cctgatggc aggccgtcca 9480
 cgaagcgttc gatgatctt cgcggcgcc ggcgcattgtt ttccgtgacg ggcggccgt 9540
 tctctcgaaa cgcagttcg aagacgcccc cgcgcaggcc gcccggccg ccgagagggg 9600
 gcaggaggcg gggcccttcg ggcagcgaga gggcgctgac gatgcaccgt atcatcttt 9660
 gcgttaggtac agctctccag gagtcgttga gcgagtccag ttggacggga tccgagaact 9720
 ttccgaggaa agcttcgatc caatcgact cgcaaggtaa gctgaggacg gtggatgag 9780
 gggcttggcg ggaggcgag gcccgcagaag aggaggaggaa gggcaggctg gaggtgatgc 9840
 tgctgatgat gtaattgaag taggcggttt tcaaaccggcg gatggtggcg aggaggacga 9900
 cgtctttggg cccggccctgc tggatgcgc ggcggtcggc catgccccag gctggctct 9960
 ggcacggcg caggccctt tagtagtctt gcatgagtct ctcgacgggg acgtcgctt 10020
 cgtcgcccg gtcggccatg cgggtggagc cgaacccgcg cagggctgc agcaggcc 10080
 ggtcgccgac cacgcgttcg gccagcacgg cctgctggat ctgggtgagc gtggcttgg 10140
 agtcgtccag gtccacgaag cgggtggtagg agccctgttt gatggttagt gtgcagttgg 10200
 ccatgacgga ccagttgacg acttgcattc cgggtgggt gatctcggtg tagcggaggc 10260
 gcgagtaggc ccgcgactcg aagacgtagt cttgcaggt ggcacgagg tactggtagc 10320
 cgacgaggaa gtgcggcgcc ggctcgccgt agagggccg ggcacggcg gccccggcgc 10380
 cggggccag gtcctccagc atgaggcggt ggttagtggta gacgtacgc gagagccagg 10440
 ttagtgcggc ggcggagggt ggcggcgccg cgaagtcgcg gacgcggctc cagatgttgc 10500
 gcaagggggc gaagcgctcc atggggca cgctctggcc ggtgaggccg ggcagtcct 10560
 gcacgctcta gacgggacag agagcgggag gttagcggct cgcgtccgtg gctggggga 10620
 cagaccgcca gggtgcgacg gggggaaacc cgggttcgag accggctgga tccgtccgtc 10680
 cccgacgcgcg cggccccgcg tccacgaccc caccagaggc cgagacccag cgcgggtgcc 10740

cggaccccaag atacggaggg gagcctttt gtggttttt cccgttagatg catccggtgt 10800
 tgcgacagat gcgccccgtcg ccagcgccgc cgacgcagcc gcccgtcccg ccccccacta 10860
 gccgcgcgga ggctctgtcc ggccggccgcg gcaaccggga ggaggaggcc atcctcgact 10920
 tggaagaagg cgagggcctg gcccggctgg gagcgcgcctc .ccccgagcgc catccccgca 10980
 tgcagctggc gagagactcg cgccaggcc acgtgccgc gcagaatctg ttcaggacc 11040
 gcagcggcca ggagccccgag gagatgaggg accgcagggtt tcacgcgggg cgggagctgc 11100
 gccgcgggctt cgaccgtcg gggtgttgc gcccgaaga cttcgagccc gacgagcgc 11160
 gcccggtaag tccggcacgg gcccacgtgt cggccggccaa cctggtgacc gcttacgagc 11220
 agacggtgaa cgaggagcgg agctttcaga aaagcttcaa caaccacgtg cgccacctga 11280
 tgcgcgcga ggaggtggcc atcggcctga tgcacatgtg ggactttgtg gaggcgtacg 11340
 tgcagaaccc gtcgagcaag ccgctgacgg cgccatgtt cctgatcgtg cagcacagtc 11400
 gggacaacga gacgttccgc gaggcgatgc tgaacatcgc ggagccccgag ggccgctggc 11460
 tcttggacct gattaacatc ctgcagagca tctgttgca ggagcgcgc ctgagcctgg 11520
 ccgacaaggt ggcggccatc aactacagca ttttggccct gggcaagttt tacgcccgc 11580
 agatctacaa gagccccctac gtgccccatag acaaggaggt gaagatcgc acgttttaca 11640
 tgcggatggc gctgaaaatgt ctgcacgtga gcccacgtatc ggggggtgtac cgcaacgc 11700
 gcatccacaa ggccgtgagc gcccggcc ggcgcgagct gagcgcaccgc gagctgatgc 11760
 acagcctgcg gagggcgctg gcggggcgccg gcccggcgca ggaggcccgag ttctacttgc 11820
 acatgggggc ggacttgcag tggcagccca gcccggccg cctggaggcg gcccgtacc 11880
 gcccggccgg cggcggtggc gagggcgagg acgaggacga ggtggagtac gaggaggagg 11940
 actgatcgcc gaggtgtttt cgttagatgca gcccggccgacg gcccggcgca gcccggccgca 12000
 gggggacccc gcccgtgtgg cggccctgca gagccaaacct tccggcgtga acgcctccgc 12060
 tgactggcg gcccggatgg accgcatttt ggccttgcacc acccgcaacc cccgaggccct 12120
 tagacagcag cccgcaggcca accgcattttc gcccatttg gaagccgtgg tccctcgcg 12180
 caccaacccc acgcacgaga aggtccctggc ggtggtaac gcccgtgtgg agagcaaggc 12240
 gatccgcaag gacgaggcg ggctgattta caacgcctg ctggagcggg tggcgcgcta 12300
 caacagcacc aacgtgcagg ccaacctgga ccgtctgacg acggacgtgc gggaggcggt 12360
 ggcgcagcgg gagcgcttca tgcgcgacac gaacctggc tgcaggtgg ccctgaacgc 12420
 cttcctgagc aegcagccgg ccaacgtgcc ggcgcggcag gaggactacg tcaagttcat 12480

cagcgcgctg cgccctcctgg tggccgaggt gccgcagac gaggtgtacc agtcgggtcc 12540
 ggactacttc ttccagacct cgcgccaggc cctgcagacg gtgaacctga cgcaggcctt 12600
 caagaacctg gaaggcatgt ggggcgtgcg ggccccgtg ggccgaccggg cgacgatctc 12660
 cagttgctg acgcccaca cgcggctgtc gctgtgtcg atcgccctt tcaccaatag 12720
 cagtaccatc agccgcgact cgtacctggg ccacctgate acgctgtacc gcgaggccat 12780
 cgggcaggcg caggtggacg agcagacctt ccaggagatt acgagcgtga gccggccct 12840
 ggggcagcag gacacgggta gcctggaggc gacgctgaat ttctgctga ccaaccggcg 12900
 gcagaagatc ccctcccagt acacgctgag cacggaggag gagcgcacatct tgctacgt 12960
 gcagcagtcc gtgagcctgt atctgatgcg cgagggggcg agcccctcg cggcgttgaa 13020
 catgacggcc cgtaacatgg agccgtcgct gtacgcccgc caccggccgt tcgtgaaccg 13080
 cctgatggac tacctgcaacc ggcgcgcgc catgaacggc gagtacttta cgaacgccc 13140
 cctgaaccccg cactggatgc cggcgtccgg ttctcacacg gggactttg acatgcccga 13200
 gggcgcacgac gggttcctgt gggacgacgt gtcggacagc gtgttcgcgc cggtgcgtcc 13260
 gggcaagaag gagggcggcg acgagctgcc gctgtccgtg gtggaggcg cgtcgcgcgg 13320
 ccagagcccg ttccccagcc tcccgctgtt gtcggcggc acgagcagcg gccgggttc 13380
 ggcgcgcgg ctggagggcg actacctgaa cgaccgctg ctgcgcggc cccggcccaa 13440
 gaactttccc aacaacgggg tggagagcct agtggataag atgaatcgct ggaagaccta 13500
 cgcgcaggag cagcgggagt gggaggagag tcagccccgc cccctgcctc cgccgcgtcc 13560
 caggtggcgc cggcgaaaag aagacccgga agactcgccg gacgatagca gctgttttgg 13620
 cttggggggg accggtgccg ctcgacaaa cccgttcgcc cacctgcgcc cgcaaggccg 13680
 gctgggtcgg ctgtatttag gaaagaaaact aataaaagaa aaaagagctt gcttaccaga 13740
 gccatggtcg cagcgtcggtt cccttgcgt gtgtttctc ctccccggta gcgaa atg 13798
 Met
 1
 agg cgc gcg gtg gga gtg ccg ccg gtg atg gcg tac gcc gag ggt cct 13846
 Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly Pro
 5 10 15
 cct cct tct tac gaa acg gtg atg ggc gcc gcg gat tcg ccg gcc acg 13894
 Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala Thr
 20 25 30
 ctg gag gcg ctc tac gtc cct ccc cgc tac ctg ggg cct acg gag ggg 13942
 Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu Gly

35

40

45

agg aac agc atc cgt tac tca gag ctg gcg ccg ctg tac gac acc acc Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr Thr 50 55 60 65	13990
cgc gtg tac ctg gtg gat aac aag tcg gcg gac atc gcg tcg ctg aac Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Ile Ala Ser Leu Asn 70 75 80	14038
tac cag aac gac cat agc aac ttt ctg acc acg gtg gtg cag aac aat Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn Asn 85 90 95	14086
gac ttt acc ccg gtg gag gcg ggc acg cag acc ata aat ttc gac gag Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp Glu 100 105 110	14134
cgc tcg cgg tgg ggc gac ctg aaa acc atc ctg cgc acc aac atg Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn Met 115 120 125	14182
ccc aac atc aac gag ttc atg tcc acc aac aag ttc agg gcc cgg ttg Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg Leu 130 135 140 145	14230
atg gta gag aaa gtg aac aag gaa acc aat gcc cct cga tac gag tgg Met Val Glu Lys Val Asn Lys Glu Thr Asn Ala Pro Arg Tyr Glu Trp 150 155 160	14278
ttt gag ttc acc ctg ccc gag ggc aac tac tcg gag acc atg acc ata Phe Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser Glu Thr Met Thr Ile 165 170 175	14326
gac ctg atg aat aac gcg atc gtg gac aac tac ttg gaa gtg ggg cgg Asp Leu Met Asn Asn Ala Ile Val Asp Asn Tyr Leu Glu Val Gly Arg 180 185 190	14374
cag aac ggg gtg ctg gag agc gac atc ggg gtg aag ttt gac acg cgc Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr Arg 195 200 205	14422
aac ttc cgg ctg ggc tgg gac ccg gtc acc aag ctg gtc atg ccc ggc Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro Gly 210 215 220 225	14470
gtg tac acc aac gag gcc ttc cac ccc gac atc gtc ctg ctg ccc ggc Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro Gly 230 235 240	14518
tgc ggc gtg gac ttc acg cag agc cgg ctg agc aac ctg ctg ggg atc Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly Ile 245 250 255	14566
cgc aag cgg atg ccc ttc cag gcg ggt ttt cag atc atg tac gag gac Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu Asp 260 265 270	14614

ctg gag ggc ggc aac atc ccc gcc ttg cta gac gtg gcg aaa tac gag Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr Glu 275 280 285	14662
gcc agc att cag aag gcg cgg gag cag ggc gag atc cgc ggc gac Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Glu Ile Arg Gly Asp 290 295 300 305	14710
aac ttt acc gtc atc ccc cgg gac gtg gag atc gtg ccc gtg gag aag Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu Lys 310 315 320	14758
gat agc aag gac cgc agt tac aac cta ctc ccc ggc gac cag acc aac Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr Asn 325 330 335	14806
acg gcc tac cgc agc tgg ttc ctg gcc tac aac tac ggc gac ccc gag Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro Glu 340 345 350	14854
aag ggc gtc agg tcc tgg acg ctg ctg acc acc acg gac gtc acc tgc Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Asp Val Thr Cys 355 360 365	14902
ggc tcg cag cag gtg tac tgg tcg ctc ccg gac atg atg caa gac ccc Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp Pro 370 375 380 385	14950
gtg acc ttc cgg ccc tcc agc caa gtc agc aac tac ccc gtg gtg gga Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val Gly 390 395 400	14998
gtc gag ctc ctg ccg gtg cac gcc aag agc ttt tac aac gag cag gcc Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln Ala 405 410 415	15046
gtc tac tcg cag ctc atc cgc cag tcc acc gcg ctc acg cac gtc ttc Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val Phe 420 425 430	15094
aac cgc ttc ccc gag aac cag atc ctg gtg cgc ccg ccc gct ccg acc Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro Thr 435 440 445	15142
att acc acc gtc agt gaa aac gtt ccc gcc ctc aca gat cac gga acc Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly Thr 450 455 460 465	15190
ctg ccg ctg cgc agc agt atc agt gga gtc cag cgc gtg acc atc act Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile Thr 470 475 480	15238
gac gcc cgg cga agg acc tgc ccc tac gtg cac aag gcc ctg ggc ata Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly Ile 485 490 495	15286

gtc gct ccc aaa gtg ctc tct agc cgc acc ttt taa caagcatgtc Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe	15332
500	505
catttcatac tcgccccaca acaacacccgg ctggggcctg cgctcgcccg gcatgtacgg cggcgccaag cgccgcgtcca gcgagcaccc cgccgcgtc cgccgcact accggggcccc	15392
ctggggcgcg cacaaggcgcg gcgtctccac ggcaccacc gtcgacgacg ccatcgacgc	15452
cgtcgtggcc caggccagac gctaccgccc gcccaagtcg acggtgacg ccgtcatcga	15512
cagcgtggtg gccgacgcgc ggccgatacgc tcgacgcaag cggcgctcgc accggcgctcg	15572
ccgtcccacc gccgccatgc tggccgcccag agccgtcctg agacgcgcgc gccgcgtggg	15632
ccgcccagcc atgcgccgag ccgcggccaa cgccagcgcg ggtcgcgcgc gtgtcaggg	15692
cgcggcag gccgcccgcg ccatcgcacaa cctggcccaa cccgcgcgg gaaacgtgta	15752
ctgggtgcga gacgcgtcgg gcgtgcgcgt gccggcgcgc acccgcccccc ctccggagtt	15812
gaagacaaaa agacggacga agactgagtt tccctgtcgt tgccagcatg agcaagcgca	15872
agttcaaaga agagctgctg gaggccctcg tgcccgagat ctacggcccg gccggcgctg	15932
ccgcccgggt ggcggacgtc aagccgaaag ttaagcccccg cgctgtgaag cgggttaaaa	15992
agcgggaaaa gaaagaggag aagcaggaag cagggttgcgt agacgtcgac gacggcgtgg	16052
agttcgtgcg gtcccttcgcg ccccgctcgcc gggtgcagtgcgcgttgcgcgtcaagc	16112
tgcgtcccgccg gccggggcacc gtgggtgttt tcaccccccgg cctgcgttcgcg gccacgcgcg	16172
gcctgaagcg cgagtacgcac gagggttatgcgac gacgcgtacgc gacgcgtacgc	16232
agcagctcgg ggagtttgcgt tacggcaagc gcccgcgtca cggggaggtgcgcgtggcgc	16292
tggaccaggg caatcccacg cccagcctca agccgcgtacgc gctgcgtcccg	16352
tgagcgcgtc gaccgagagc aagcggggca tcaagaggaa gatggggcgcac ctgcagccca	16412
ccatgcaact catggtgccc aaacggcaga agctggagga cgtgtggag aacatgaaag	16472
tggatcccacg catcgagccc gaagtgaaag tgcgcgtccat caagggatgcgcgtggcc	16532
taggcgtgcgtca gacgggtggac attcagatcc cccgcgtccgc ctcccccgtt tctgcacca	16592
ctacgacggc cgtggaggcc atggaaacgc agacggagct gcccgcggcc ttggcggcag	16652
ccgcccaccgc cgccgcggct acccgagaga tgggcgtgcgtca gaccgacccc tggtacgagt	16712
tgcgtccggcccg cgccgcgtcgccat ccacgagccc gtcggatgcgcgtccgc ggcgaccacc	16772
ctgactacgt cttgcgtccct tccatcgcacgc cgcacgcggcc ctacggcggaa acgacccgttcc	16832
gccccgggtcg cgccgcgtccgc accaccgcgc gtcgtcgac cccgcgcgc cgtcgagcc	16892
	16952

gtcgcgcact ggctccatc gcggttcgcc gctcgatcg ccggggtcgc acgctgaccc 17012
 tgcccaccgc gcgttaccac cccagcatcg tcatttaacc tgcgctgccc tttgcagat 17072
 ggctctgacg tgccgtttc gttccccgt tcggactac cgaggaagat ctgcggtag 17132
 gactggtcta gcggcagcg gtctccgacg ccggccgcgc gcggtgcacc ggcgcatgaa 17192
 gggcggcatt ctgcccgcgc tgatccccat tatcgccgcc gccatgggg cgatccccgg 17252
 cgtggcctcg gtggccttgc aagcagctcg caaaaattaa ataaagaagg cttgacactc 17312
 actgccttgtt ctgactgtt tcatgcagac aagacatgga agacatcaat tttgcgtcg 17372
 tggcccccgcg gcacggctcg cggccgttca tggcacctg gaacgagatc ggcaccagcc 17432
 agctcaacgg gggcgcttgc agttggagca gcctgtggag cggcattaaa aactttgggt 17492
 ccacgattaa gacctatggc aacaaggcgt ggaacagtag cactggtcag atgcctcg 17552
 ataagctgaa ggaccagaac ttccagcaga aagtggtaga cggtctggcc tcgggcatca 17612
 acggggtgtt ggacctggcc aaccaggcgg tgcagaacca gatcaaccag cgtctggaga 17672
 acagccgcca gcegccccgcg gcccgcgcg agcgtccgca ggtggaggag gtggaaagtgg 17732
 aggagaagct gccgccccctg gagacggtgt cggccgtggg cgtgccttgc aagggggaga 17792
 agcggccgcg gcccgcgcg gaggagaccc tagtgcaccga gaccctggag cggccctcg 17852
 acgagcaggc ttgaaagag gggccacgc ccctgccccat gaccggcccc atcgacc 17912
 tggcccgacc ggtctacggc aaggaacaca aagccgtgac gctagagctg cctccgccc 17972
 cggccaccgt acccccgatg cccggccca ccctggcac cggccgtgcct cgtccccccg 18032
 ccccgccggt cggccgtggcc acgccccgcg gcccgcgcg cggagccaac tggcagagca 18092
 ctctgaacag catcggtggc ctgggagtgaa aaacccctgaa acgccccggg tggactatt 18152
 aaagccagct aaatacccat gtgttgtatg cgcctctgt gtcacgcccag aaaaagccag 18212
 cccgagtgcg ggtcaccgccc gcccggccaa ggcggccgttt caag atg gcc acc ccc 18268
 Met Ala Thr Pro
 510
 tcg atg atg ccg cag tgg tct tac atg cac atc gcc ggg cag gac gcc 18316
 Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala Gly Gln Asp Ala
 515 520 525
 tcg gag tac ctg agc ccg ggc ctg gtg cag ttc gcc cgc gcc acc gac 18364
 Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala Arg Ala Thr Asp
 530 535 540
 acg tac ttc agc ctg ggc aac aag ttt agg aac ccc acg gtg gcc ccc 18412
 Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro Thr Val Ala Pro

545	550	555	560	
acc cac gac gtg acg acg gac cgg tcc cag cgg ctg acg ctg cgg ttc Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu Thr Leu Arg Phe				18460
565	570	575		
gtg ccc gtc gac cgc gag gac acc gcg tac tcg tac aaa gtg cgc ttc Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr Lys Val Arg Phe				18508
580	585	590		
acg ctg gcc gtg ggc gac aac cgc gtg gac atg gcc agc acg tac Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met Ala Ser Thr Tyr				18556
595	600	605		
ttt gac atc cgc ggc gtg ttg gac cgc ggt ccc agc ttc aaa ccc tac Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser Phe Lys Pro Tyr				18604
610	615	620		
tcc ggc acc gcc tac aac tcc ctg gcc ccc aag ggc gcc ccc aac ccg Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly Ala Pro Asn Pro				18652
625	630	635	640	
tca gaa tgg aag ggc tca gac aac aaa att agt gta aga ggt cag gct Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val Arg Gly Gln Ala				18700
645	650	655		
ccg ttt ttt agt aca tcc att aca aag gat ggt att caa gtg gcc act Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile Gln Val Ala Thr				18748
660	665	670		
gat act tct agc gga gct gtg tat gct aaa aag gaa tat cag cct gaa Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu Tyr Gln Pro Glu				18796
675	680	685		
cca caa gta ggg caa gaa caa tgg aac agc gaa gcc agt gat agt gat Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala Ser Asp Ser Asp				18844
690	695	700		
aaa gta gct ggt agg att cta aaa gac aca aca ccc atg ttc cct tgt Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro Met Phe Pro Cys				18892
705	710	715	720	
tac ggt tcc tac gcc aag ccc aca aat gaa cag ggg ggg caa ggc act Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly Gln Gly Thr				18940
725	730	735		
aat act gta gat ctg cag ttc ttt gcc tct tca tcg gct acc tct acg Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ala Thr Ser Thr				18988
740	745	750		
cct aaa gcc gta ctc tat gcc gag gac gtg gca ata gaa gca cca gac Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile Glu Ala Pro Asp				19036
755	760	765		
acc cat ttg gtg tac aaa ccg gca gtt aca acc acg acc act agt tcc Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr Thr Ser Ser				19084
770	775	780		

caa gac ctg cta act cag cag gct gct ccc aac cga ccc aac tac att Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg Pro Asn Tyr Ile 785 790 795 800	19132
ggc ttc agg gat aat ttt atc ggt ctc atg tat tac aac tcc act ggc Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr Asn Ser Thr Gly 805 810 815	19180
aat atg ggt gtt ttg gca ggg caa gct tct cag cta aac gcg gtg gtt Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu Asn Ala Val Val 820 825 830	19228
gac ttg caa gac aga aac acc gag ctg tcc tac cag ctc atg ctt gat Asp Leu Gln Asp Arg Asn Thr Glu Leu Ser Tyr Gln Leu Met Leu Asp 835 840 845	19276
gct ttg ggc gac cgc agt cgt tac ttc tcc atg tgg aac cag gcc gta Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp Asn Gln Ala Val 850 855 860	19324
gac agc tat gac cct gat gtc aga att att gaa aat cat ggt gtg gag Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Val Glu 865 870 875 880	19372
gat gag ctg cca aac tac tgt ttc ccg cta gga ggg tcg cta gta act Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ser Leu Val Thr 885 890 895	19420
gaa act tat aca ggc cta tca ccc caa aac gga agt aac acg tgg aca Glu Thr Tyr Thr Gly Leu Ser Pro Gln Asn Gly Ser Asn Thr Trp Thr 900 905 910	19468
acc gac agc acc acc tat gca act aga ggg gtg gaa atc ggc tct ggc Thr Asp Ser Thr Thr Tyr Ala Thr Arg Gly Val Glu Ile Gly Ser Gly 915 920 925	19516
aac atg ttc gcc atg gaa att aat ttg gcg gcc aat cta tgg agg agt Asn Met Phe Ala Met Glu Ile Asn Leu Ala Ala Asn Leu Trp Arg Ser 930 935 940	19564
tcc ctg tac tcc aac gtg gcc ctg tac ctg ccc gac gag tac aag ctc Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Glu Tyr Lys Leu 945 950 955 960	19612
acc ccc gac aac atc acc ctc ccc gac aac aaa aac act tac gac tac Thr Pro Asp Asn Ile Thr Leu Pro Asp Asn Lys Asn Thr Tyr Asp Tyr 965 970 975	19660
atg aac ggc cgc gtg gcc ccc agc tcc ctc gac acc tac gtc aac Met Asn Gly Arg Val Ala Ala Pro Ser Ser Leu Asp Thr Tyr Val Asn 980 985 990	19708
atc ggg gcg cgc tgg tcc ccc gac ccc atg gac aac gtc aac ccc ttc Ile Gly Ala Arg Trp Ser Pro Asp Pro Met Asp Asn Val Asn Pro Phe 995 1000 1005	19756

aac cac cac cgc aac gcg gga	ctg cgc tac cgc tcc	atg ctg ctg	19801
Asn His His Arg Asn Ala Gly	Leu Arg Tyr Arg Ser	Met Leu Leu	
1010	1015	1020	
ggc aac ggc cgc tac gta ccc	ttc cac atc caa gtg	ccc cag aaa	19846
Gly Asn Gly Arg Tyr Val Pro	Phe His Ile Gln Val	Pro Gln Lys	
1025	1030	1035	
tac ttc gcc atc aaa aac ctc	ctg ctc ctc ccc ggg	tcc tac acc	19891
Phe Phe Ala Ile Lys Asn Leu	Leu Leu Leu Pro Gly	Ser Tyr Thr	
1040	1045	1050	
tac gag tgg aac ttc cgc aag	gac gtc aac atg atc	ctc cag agc	19936
Tyr Glu Trp Asn Phe Arg Lys	Asp Val Asn Met Ile	Leu Gln Ser	
1055	1060	1065	
agc ctg ggt aac gac ctc cgc	gtc gac ggg gcc agc	gtc agg ttc	19981
Ser Leu Gly Asn Asp Leu Arg	Val Asp Gly Ala Ser	Val Arg Phe	
1070	1075	1080	
gac agc atc aac ctg tac gcc	aac ttc ttc ccc atg	gcc cac aac	20026
Asp Ser Ile Asn Leu Tyr Ala	Asn Phe Phe Pro Met	Ala His Asn	
1085	1090	1095	
acc gcc tcc acg ctc gag gcc	atg ctg cgc aac gac	acc aac gac	20071
Thr Ala Ser Thr Leu Glu Ala	Met Leu Arg Asn Asp	Thr Asn Asp	
1100	1105	1110	
cag tcg ttc aac gac tac ctc	tgc gct gcc aac atg	ctc tac ccc	20116
Gln Ser Phe Asn Asp Tyr Leu	Cys Ala Ala Asn Met	Leu Tyr Pro	
1115	1120	1125	
atc ccc gcc aac gcc acc agc	gtg ccc atc tcc att	ccc tcg cgg	20161
Ile Pro Ala Asn Ala Thr Ser	Val Pro Ile Ser Ile	Pro Ser Arg	
1130	1135	1140	
aac tgg gcc gcc ttc cgg ggc	tgg agc ttc acc cgg	ctc aag acc	20206
Asn Trp Ala Ala Phe Arg Gly	Trp Ser Phe Thr Arg	Leu Lys Thr	
1145	1150	1155	
aag gag acc ccc tct ctg ggc	tcc ggc ttc gat ccc	tac ttc acc	20251
Lys Glu Thr Pro Ser Leu Gly	Ser Gly Phe Asp Pro	Tyr Phe Thr	
1160	1165	1170	
tac tcg ggc tcc atc ccc tac	ctg gac ggc acc ttc	tac ctc aac	20296
Tyr Ser Gly Ser Ile Pro Tyr	Leu Asp Gly Thr Phe	Tyr Leu Asn	
1175	1180	1185	
cac act ttc aag aag gtc tcc	atc atg ttc gac tcc	tcc gtc agc	20341
Mis Thr Phe Lys Lys Val Ser	Ile Met Phe Asp Ser	Ser Val Ser	
1190	1195	1200	
tgg ccc ggc aac gac cgc ctg	ctg acc ccc aac gag	ttc gag atc	20386
Trp Pro Gly Asn Asp Arg Leu	Leu Thr Pro Asn Glu	Phe Glu Ile	
1205	1210	1215	
aag cgc acc gtg gac ggg gaa	ggg tac aac gtg gcc	cag tgc aac	20431

Lys	Arg	Thr	Val	Asp	Gly	Glu	Gly	Tyr	Asn	Val	Ala	Gln	Cys	Asn		
1220						1225					1230					
atg	acc	aag	gac	tgg	ttc	ctc	atc	cag	atg	ctc	agc	cac	tac	aac	20476	
Met	Thr	Lys	Asp	Trp	Phe	Leu	Ile	Gln	Met	Leu	Ser	His	Tyr	Asn		
1235						1240					1245					
atc	ggc	tac	cag	ggc	ttc	tac	gtg	ccc	gag	ggc	tac	aag	gac	agg	20521	
Ile	Gly	Tyr	Gln	Gly	Phe	Tyr	Val	Pro	Glu	Gly	Tyr	Lys	Asp	Arg		
1250						1255					1260					
atg	tac	tct	ttc	ttc	cgc	aac	ttc	caa	ccc	atg	agc	cgc	cag	gtg	20566	
Met	Tyr	Ser	Phe	Phe	Arg	Asn	Phe	Gln	Pro	Met	Ser	Arg	Gln	Val		
1265						1270					1275					
gtc	gac	acc	acc	acc	tac	acc	gac	tac	aaa	aac	gtc	acc	ctc	ccc	20611	
Val	Asp	Thr	Thr	Thr	Tyr	Thr	Asp	Tyr	Lys	Asn	Val	Thr	Leu	Pro		
1280						1285					1290					
ttc	cag	cac	aat	aa	tcg	ggg	ttc	gtg	gga	tac	atg	ggc	ccc	acc	20656	
Phe	Gln	His	Asn	Asn	Ser	Gly	Phe	Val	Gly	Tyr	Met	Gly	Pro	Thr		
1295						1300					1305					
atg	cgc	gag	ggg	cag	gcc	tac	ccc	gcc	aac	tac	ccc	tac	ccc	ctg	20701	
Met	Arg	Glu	Gly	Gln	Ala	Tyr	Pro	Ala	Asn	Tyr	Pro	Tyr	Pro	Leu		
1310						1315					1320					
atc	ggc	aag	acc	gcc	gtg	ccc	agc	ctc	acg	cag	aaa	aag	tcc	ctc	20746	
Ile	Gly	Lys	Thr	Ala	Val	Pro	Ser	Leu	Thr	Gln	Lys	Lys	Phe	Leu		
1325						1330					1335					
tgc	gac	cgc	acc	atg	tgg	cgc	atc	ccc	ttc	tcc	agt	aac	tcc	atg	20791	
Cys	Asp	Arg	Thr	Met	Trp	Arg	Ile	Pro	Phe	Ser	Ser	Asn	Phe	Met		
1340						1345					1350					
tcc	atg	ggg	gcg	ctc	acc	gac	ctg	ggg	cag	aac	atg	ctg	tac	gcc	20836	
Ser	Met	Gly	Ala	Leu	Thr	Asp	Leu	Gly	Gln	Asn	Met	Leu	Tyr	Ala		
1355						1360					1365					
aac	tcc	gcc	cac	gcc	ctc	gac	atg	acc	ttc	gag	gtg	gac	ccc	atg	20881	
Asn	Ser	Ala	His	Ala	Leu	Asp	Met	Thr	Phe	Glu	Val	Asp	Pro	Met		
1370						1375					1380					
gat	gag	ccc	acg	ctt	ctc	tat	gtt	ctg	ttc	gaa	gtg	ttc	gac	gtc	20926	
Asp	Glu	Pro	Thr	Leu	Leu	Tyr	Val	Leu	Phe	Glu	Val	Phe	Asp	Val		
1385						1390					1395					
gtg	cgc	atc	cac	cag	ccg	cac	cgc	ggc	gtc	atc	gag	gcc	gtc	tac	20971	
Val	Arg	Ile	His	Gln	Pro	His	Arg	Gly	Val	Ile	Glu	Ala	Val	Tyr		
1400						1405					1410					
ctg	cgc	acg	ccg	ttc	tcg	gcc	ggt	aac	gcc	acc	acc	taa	ggagggggcc		21020	
Leu	Arg	Thr	Pro	Phe	Ser	Ala	Gly	Asn	Ala	Thr	Thr					
1415						1420					1425					
ggccgacggat gggctccagc gagccggagc tggtcgccat cgcgccgac ctgggctgcg																21080

ggccctactt cctgggcacc tttgacaaac gcttcccggg cttcggtggc cgcacaagc 21140
 tggcctgcgc catcgtaac accgcccggac gcgagaccgg cggcgccac tggctggccc 21200
 tggcctggaa ccccccgagc cgaacctgct acctcttca ccccttcggc ttctcggaacg 21260
 acaggjtcag gcagatctac cagttcgagt acgaaggcct gctccggcgc agcgccctcg 21320
 cctccacccc cgaccactgc gtacccctcg tcaagtccac ccagaccgtc caggggcccc 21380
 gctcgccgc ctgcggcctc ttctgctgca tggtcctgca cgccttcgtc cgctggcccg 21440
 cctcccccat ggacggcaac cccaccatgg acctccttac gggcggtccc aacagcatgc 21500
 ttcagagtc ccaagggtcgag cccaccctcc accgcaacca ggaggaactc tacgccttcc 21560
 tggctcgga ctcggctac ttgcggcc accgctgagcg catagaaaag gccaccgcgt 21620
 ttgacaaaat gaacgactag attttctgtg aaaaacactc aataaagcct ttattggttc 21680
 accacacgtg caegcatgca gacttttat taaaagggc tccgccttc cgtcgccgtg 21740
 gctgggtgggg agggagacgt tgcgatactg caggcgaggag ctccatctga actcgaaat 21800
 cagcagcttg ggcagggggc ctcgacgtt ctcgtccac agcttgcgca ccagctgcag 21860
 ggcgcggcaggc aggtcgccgc cggagatctt gaagtcgcag ttggggccct ggttgcgcgc 21920
 ggagttgcgg tacaccgggt tggcgactg gaacaccagc acgctgggggt gctcgatgtc 21980
 ggcgcggcc gtcttgcgg tcacctcgcc ggcgcgcagg gactccgcgt tgctcagcgc 22040
 gaaggcggtc agcttgcaca gctgccgacc cagcacgggc accccgcgtc gctggttcag 22100
 gcagtcgcag cgcatacgcca tcagcagccg cttctgcccc tgctgcatct tcggatagtc 22160
 ggctcgcatg aaggcctcca tctgcccggaa ggccgtctgc gccttgcgtc cctccgagaa 22220
 gaacagcccg caggacttgc cggagaacac gttgtgcgg cagctacgt cttccacgca 22280
 gcagcgcgcg tcgtcggttc tcaagtcgac cacgctgcgg ccccagcggt tctgcaccac 22340
 ctgggtcttgc cgggatgtt cttcagggc ctcgtggccg ttctcgctgg tcacgtccat 22400
 ctccaccacc tgctccttc ggatcatctc cagcccggtgg tagcagcgca gcacgcctc 22460
 ctgcgtcggtg caccctgtca ggcagacggc gcagccggc ggctccagct gttgaggttt 22520
 caccggcgcg taggtctcca cgtacggccg caggaagcgg cccatcatct ccacaaaggt 22580
 cttctgaccg gtgaagggtca gctgcagccg gcgtatgtcc tctgtgagcc acgtctgaca 22640
 gatcttgcgg tacaccattgc cctgtcggtt cagaaacttg aaagcggccct tctctcggtt 22700
 ctccacgtgg tacttctcca tcagcgcga catcagctcc atgccttc cccaggccga 22760
 caccagcgcc tccgcgcggg ggttaccac cgccatgcct cggaaagtgc cggggcgctc 22820

atcttccctcc tccctccctgt cttcttcttg aggccggcggt ggcggcagtt gtctcacgaa 22880
 tctcttgcgg ttggccttct ggacgatctc cacgcggggg tgggtgaacc cgtggggcac 22940
 caccacttcg tcccttcttct cttcgctgtc gggcacgact tcgggagagg gagggcgccgg 23000
 aggaaccgggt gcggccactg cggccatcgc ggcgttcttg cgcccttct tggggggcag 23060
 aggccggcgtc tccgttcccg ggctggcttc ttgcaggttag ggcgtgtatgg tgtgggaggt 23120
 ggggcgtctt ctgcgtacggc cggccatgtct gatgcgttgc tccctaggcga aaagatggag 23180
 gaggatctta gacagccgca gcccgtctcc gaaaccttaa ccacccccgc ctctgaggc 23240
 ggcgcggcg agctagacat gcaacgggag gaggaggagg acgtgcgagt ggagcaagac 23300
 ccgggctacg tgacgcccgc cgaggacggc gagggagccgc aggacccggc gccaacgctc 23360
 agcgaagccg actacctggg aggggaggac gacgtgtcgc tgaagcacct ggccggcag 23420
 agcaccatcg tgcaggaggc cctcaaggag cgcgaggagg tcccgctgac ggtggaggag 23480
 ctcaagccggg cctacgaagc caaccttctc tcgcccggg tgcccccaa gaagcaggcc 23540
 aacggcacct gcgagcccaa ccccccctc aacttctacc cctgtttgc ggtggccgag 23600
 gegctggcca cctatcacat cttcttcaag aaccagcgca tccccctctc gtgccgcgc 23660
 aaccgcaccc gcgcgcaccc cctccctgcat ctcccgagccg gcgcgcaccc acctgagatc 23720
 gcctccctgg aggaagtccc caagatcttca gaaaggctcg gcaaggacga gaagcgcgcg 23780
 gcaaacgctc tggaaaagaa cgagagcgag ggtcagaacg tgctggcga gctggaaggc 23840
 gacaacgcgc gtcgtggcggt gctcaaacgc accatcgaaag tctccctactt cgcctacccc 23900
 ggcgtcaacc tcccccccaa ggtcatgcgc tcggcatgg atcagctgtct catcaagcgc 23960
 gccgagccccc tcgagaacga ctcccgagggt gattccgagg acggaaaaacc cgtggctcg 24020
 gacgaggaggc tcgcgcgtcg gctggcaccg caggaccccg ccgagttgca agagcggcgc 24080
 aagatgtga tggcgcccggt gctggtcacc gcccggcgtcg agtgcgtcga ggccttctc 24140
 gccgacccccc agaccctcgcaaggctcgag gagtccctgc actacgcctt cgcgcacggc 24200
 tacgtgcgc accgcctgcac gatctccaac gtggagctta gcaacctgggt ctccctacatg 24260
 ggcatactgc acgagaacccg cctcggcgag aacgtctcc actgcacccct gaccggggag 24320
 gcccgcgcg actacgtcccg cgactgcac taccttttc tccacccctac ctggcagacc 24380
 gccatggggg tctggcageca gtgtctggag gagcgcaacc ttcgcgagct cgacaagcta 24440
 ctgagccgcg agcgccgcga gctctggacg gctttcagcg agcgcaccgc cgcctggcg 24500

ctggccgacc tcataatccc cgagcgactc aggcaaacc ctcagaacgg cctgccccac 24560
 tttgtcagcc agagcatgt gcaaaaactt cgttcattca tcctggagcg atccggatc 24620
 ttggccgcca tgagctgcgc cctgcctcc gatttcgtcc ccctctatta tcgcgagtgc 24680
 ccccccgcgc tctggagcca ctgcgtacctg ctgcgtctgg ccaactacct cgccccaccac 24740
 tccgacctca tggaagactc cagcggcgag gggctgctgg agtgcactg cgcgtcaac 24800
 ctctgcaccc cccacccgctc gctggctgtc aacaccgagc tgctcagcga gacgcaagtg 24860
 atcggtaacct ttgagatcca gggaccagag gggccggagg gtgcattccaa cctcaagctc 24920
 agcccccggcgc tctggacttc cgcctacctg cgcaaatata tccccgagga ctatcacgcc 24980
 caccagatcc aattctacga agaccaatcg cgacccccc aagccccct cacggctgt 25040
 gtcatcaccc agagccagat tctggccaa ttgcaagcca tccagcaggc cccccaagag 25100
 ttccctcctga aaaagggtca cggggcttat ctggacccccc agaccggcga ggaactcaac 25160
 accccgtcac cctccggcgc cgcttcgtgc cgcccgacga accatgccgc ccaaagggaa 25220
 caagcaggcc atcgcccagc ggccggccaa gaagcagcaa gagctccagg agcagtggga 25280
 cgaggagtcc tgggacagcc aggcggagga agtctcagac gaggaggagg acatggagag 25340
 ctgggacagc ctagacgagg aggaggaggc cgaggagcta gaggacgagc ctctcgagga 25400
 ggaagagccc agcagcgccg cggcaccatc ggcttccaaa gaagcggctc ggagccggcc 25460
 ggccccgaag cagcagaagc agcaacagcc gccaccgtcg ccccccgcgc caccaccagg 25520
 ctcactcaaa gccagccgt a ggtggacgc ggtgtccatc gcgggatcgc ccaaagcccc 25580
 agtcggtaag ccacccgggc ggtcgccgc ggggtactgt tcctggcgcc cccacaagag 25640
 caagatcgcc gctgcctcc agcaactgcgc gggcaacatc tccttcgtgc ggcgtactt 25700
 gtcattccac gacgggggtgg cggtgccgcg caacgtctc tactattacc gtcatctata 25760
 cagccctac gagacagaag gcccggcctc cgcgtaaagac cagccgccag acggtctct 25820
 ccgcacatgc gacccggccag gactcggccg ccacgcagga gtcagaaaa cgcacatccc 25880
 ccacccgtta tgctatcttc cagcagagcc gcccggccagca gtcggaaactg aaagtaaaaa 25940
 accgcctccct gcgttcgtctc acccgagct gtcgtacca caggaggaa gaccaactgc 26000
 agcgcacgct cgaggacgccc gaggactgt tcaataaata ctgctcggtg tctcttaagg 26060
 actgaaagcc cgcgtttttt cagaggctca ttacgtcatc atcatcatga gcaaggacat 26120
 tccccacgcct tacatgtgga gtcaccagcc gcaatggga ctggccggccg ggcctccca 26180
 ggattactcc agtcgcacatga actggctgag tgccggccccc cacatgatcg ggcgggtcaa 26240

tgggattcgt gccacccgca atcagatact gctggaacag gcccgcctca cctccacccc 26300
 gcgacgtcag ctgaacccgc ccgcgtggcc cgccgcccag gtgtaccagg aaaaccccg 26360
 cccgaccaca gtccctcctgc cacgcacgc ggagggcgaat gtcagatga ctaactccgg 26420
 ggcgcaatta gccccggcg cccgacgt cgtcgctccc gggtacagag gtcggcccg 26480
 accctacccc tccggcccta taaagaggct gatcattcga ggccgaggta tccagctcaa 26540
 cgacgaggtg gtgagctct cgaccggctc tcggcccgac ggagtcttcc agcttggagg 26600
 cgcggccgc tcttccttca ccactcgcca ggcttacctg acgctccaga gctttccctc 26660
 ccagcctcgc tccggcggca tcggcacccct ccagttcgtg gaggagttcg tgccctcggt 26720
 ctacttcaac ccgttctccg gcttcccg cccgttacccg gacagcttca tcccaacta 26780
 cgacgcggtg agcgaatccg tggacggcta cgattgatga ccgatggtgc ggccgtaact 26840
 ggcggccgc aacatctgca tcactgccat cgttcccggt gcttcgcccggaggcctgt 26900
 gagttcatct acttccagct cgcccccggac cagcttcagg gcccggca cggcgtaag 26960
 ctcgtgatag aggaagagct cgagagtagc tgctcgct gtttacccctc ggcggccatc 27020
 ctagtcgaga gggAACGCGG taggaccacc tcaccctct actgcattcg tgactccccg 27080
 gaattacatg aagatctgtg ttgccttcta tgtgccgaac aataaccctt cttgtaaacta 27140
 cctacatcca caataaacca gaatttggaa actcccttcg tttgtttgca g atg aaa 27197
 Met Lys

cgc	gcc	cgc	ctc	gac	gac	ttc	aac	ccc	gtc	tac	ccc	tat	gac	27242	
Arg	Ala	Arg	Ile	Asp	Asp	Asp	Phe	Asn	Pro	Val	Tyr	Pro	Tyr	Asp	
1430								1435					1440		
act	ccc	aac	gct	ccc	tct	gtt	ccc	ttc	atc	act	cct	ccc	ttc	gtc	27287
Thr	Pro	Asn	Ala	Pro	Ser	Val	Pro	Phe	Ile	Thr	Pro	Pro	Phe	Val	
1445								1450					1455		
tcc	tcg	gac	ggc	ttg	caa	gaa	aaa	cca	ccc	gga	atg	ctc	agt	ctc	27332
Ser	Ser	Asp	Gly	Leu	Gln	Glu	Lys	Pro	Pro	Gly	Met	Leu	Ser	Leu	
1460								1465					1470		
aac	tac	caa	gat	cct	att	acc	acc	caa	aac	ggg	gca	tta	act	cta	27377
Asn	Tyr	Gln	Asp	Pro	Ile	Thr	Thr	Gln	Asn	Gly	Ala	Leu	Thr	Leu	
1475								1480					1485		
aag	ctt	ggc	agc	gga	ctg	aac	ata	aac	caa	gat	ggg	gaa	ctt	acc	27422
Lys	Leu	Gly	Ser	Gly	Leu	Asn	Ile	Asn	Gln	Asp	Gly	Glu	Leu	Thr	
1490								1495					1500		
tca	gac	gcc	agc	gtt	ctc	gtc	act	ccc	ccc	att	aca	aaa	gcc	aac	27467
Ser	Asp	Ala	Ser	Val	Leu	Val	Thr	Pro	Pro	Ile	Thr	Lys	Ala	Asn	

1505	1510	1515	
aac aca ata ggc cta gcc ttc aat	gca cct ctt acc ttg	caa agc	27512
Asn Thr Ile Gly Leu Ala Phe Asn	Ala Pro Leu Thr Leu	Gln Ser	
1520	1525	1530	
gat act tta aat ctt gct tgt aac	gcc cca ctt acc gtg	caa gac	27557
Asp Thr Leu Asn Leu Ala Cys Asn	Ala Pro Leu Thr Val	Gln Asp	
1535	1540	1545	
aat agg ttg gga ata aca tac aac	tct ccc ctc acc ttg	caa aac	27602
Asn Arg Leu Gly Ile Thr Tyr Asn	Ser Pro Leu Thr Leu	Gln Asn	
1550	1555	1560	
agc gaa ctt gcc cta gcg gtc acc	ccg cct ctt gac act	gcc aat	27647
Ser Glu Leu Ala Leu Ala Val Thr	Pro Pro Leu Asp Thr	Ala Asn	
1565	1570	1575	
aac aca ctt gcg ctt aaa acc gcc	cgg cct ata att aca	aac tct	27692
Asn Thr Leu Ala Leu Lys Thr Ala	Arg Pro Ile Ile Thr	Asn Ser	
1580	1585	1590	
aat aac gag ctt aca ctc tcc gct	gat gct ccc cta aac	acc agc	27737
Asn Asn Glu Leu Thr Leu Ser Ala	Asp Ala Pro Leu Asn	Thr Ser	
1595	1600	1605	
acg ggt acc ctc cgc cta caa agc	gca gca cca ctg ggg	cta gtt	27782
Thr Gly Thr Leu Arg Leu Gln Ser	Ala Ala Pro Leu Gly	Leu Val	
1610	1615	1620	
gac caa acc ctg cga gtg ctt ttt	tct aac cca ctc tac	ttg caa	27827
Asp Gln Thr Leu Arg Val Leu Phe	Ser Asn Pro Leu Tyr	Leu Gln	
1625	1630	1635	
aac aac ttt ctc tca cta gcc att	gaa cgc cca ttg gct	tta act	27872
Asn Asn Phe Leu Ser Leu Ala Ile	Glu Arg Pro Leu Ala	Leu Thr	
1640	1645	1650	
acc act ggt tct atg gct atg cag	att tcc caa cca tta	aaa gtg	27917
Thr Thr Gly Ser Met Ala Met Gln	Ile Ser Gln Pro Leu	Lys Val	
1655	1660	1665	
gaa gac gga agc tta agc ttg agc	att gaa agc cct cta	aat cta	27962
Glu Asp Gly Ser Leu Ser Leu Ser	Ile Glu Ser Pro Leu	Asn Leu	
1670	1675	1680	
aaa aac gga aat ctt act tta gga	acc caa agt ccc cta	act gtc	28007
Lys Asn Gly Asn Leu Thr Leu Gly	Thr Gln Ser Pro Leu	Thr Val	
1685	1690	1695	
act ggt aac aac ctc agc ctt aca	aca aca gcc cca tta	acg gtt	28052
Thr Gly Asn Asn Leu Ser Leu Thr	Thr Thr Ala Pro Leu	Thr Val	
1700	1705	1710	
cag aac aac gct cta gcc ctc tca	gtg tta ctg ccg ctt	aga cta	28097
Gln Asn Asn Ala Leu Ala Leu Ser	Val Leu Leu Pro Leu	Arg Leu	
1715	1720	1725	

ttt aat aac acc tca ctg gga gtg gca ttc aac cca ccc att tct		28142
Phe Asn Asn Thr Ser Leu Gly Val Ala Phe Asn Pro Pro Ile Ser		
1730 1735 1740		
tca gca aac aac ggg ctg tct ctt gac att gga aat ggc ctt aca		28187
Ser Ala Asn Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr		
1745 1750 1755		
ctg caa tac aac agg ctc gta gtg aac att ggc ggc ggg cta cag		28232
Leu Gln Tyr Asn Arg Leu Val Val Asn Ile Gly Gly Leu Gln		
1760 1765 1770		
ttt aac aac ggt gct att acc gct tcc ata aat gca gct ctg ccg		28277
Phe Asn Asn Gly Ala Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro		
1775 1780 1785		
ttg cag tat tcc aat aac cag ctt tct ctt aat att gga ggc ggg		28322
Leu Gln Tyr Ser Asn Asn Gln Leu Ser Leu Asn Ile Gly Gly Gly		
1790 1795 1800		
ctg cga tac aac ggc act tac aaa aat tta gcc gtc aaa acc gac		28367
Leu Arg Tyr Asn Gly Thr Tyr Lys Asn Leu Ala Val Lys Thr Asp		
1805 1810 1815		
tct ttt agg ggt ctt gaa att gac agt aat cag ttc ctg gtg cca		28412
Ser Phe Arg Gly Leu Glu Ile Asp Ser Asn Gln Phe Leu Val Pro		
1820 1825 1830		
aga ctg ggt tct ggt cta aag ttt gat caa tat ggg tac att agc		28457
Arg Leu Gly Ser Gly Leu Lys Phe Asp Gln Tyr Gly Tyr Ile Ser		
1835 1840 1845		
gtc ata cct cca act gtt acg cca aca aca ctt tgg act aca gca		28502
Val Ile Pro Pro Thr Val Thr Pro Thr Thr Leu Trp Thr Thr Ala		
1850 1855 1860		
gac cct tct ccc aac gct act ttt tac gac agc tta gat gct aag		28547
Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser Leu Asp Ala Lys		
1865 1870 1875		
gta tgg ctg gcc tta gta aaa tgc aac ggc atg gtt aat gga acc		28592
Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val Asn Gly Thr		
1880 1885 1890		
ata gcc ata aag gct tta aaa ggt act ctg ctc caa cct acg gct		28637
Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro Thr Ala		
1895 1900 1905		
agt ttt att tct ttt gtt atg tat ttt tac agc aat ggc acc aga		28682
Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr Arg		
1910 1915 1920		
aga act aac tac ccc acg ttt gaa aat gaa ggc ata cta gct agt		28727
Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser		
1925 1930 1935		

agt gct aca	tgg ggt tat cgt caa	gga aac tcg gca aac	acc aac	28772		
Ser Ala Thr	Trp Gly Tyr Arg Gln	Gly Asn Ser Ala Asn	Thr Asn			
1940	1945	1950				
gtc acc agt	gcc gtt gaa ttt atg	cct agc tcc aca aga	tat cct	28817		
Val Thr Ser	Ala Val Glu Phe Met	Pro Ser Ser Thr Arg	Tyr Pro			
1955	1960	1965				
gtt aac aag	ggt act gag gtt cag	aac atg gaa ctc acc	tac act	28862		
Val Asn Lys	Gly Thr Glu Val Gln	Asn Met Glu Leu Thr	Tyr Thr			
1970	1975	1980				
tcc ttg cag	gga gac ccc act atg	gcc ata tca ttt caa	gct att	28907		
Phe Leu Gln	Gly Asp Pro Thr Met	Ala Ile Ser Phe Gln	Ala Ile			
1985	1990	1995				
tat aac cat	gct ttg gaa ggt tac	tct tta aaa ttt acc	tgg cga	28952		
Tyr Asn His	Ala Leu Glu Gly Tyr	Ser Leu Lys Phe Thr	Trp Arg			
2000	2005	2010				
gtt cgc aac	agg gaa cgc ttt gat	atc ccc tgc tgt tct	ttt tct	28997		
Val Arg Asn	Arg Glu Arg Phe Asp	Ile Pro Cys Cys Ser	Phe Ser			
2015	2020	2025				
tac ata acg	gaa gaa taa acactgttt	tctttcaat	gttttattc	29045		
Tyr Ile Thr	Glu Glu					
2030						
tgctttta cacagttcga	accgtcagac	tccctcccc	cttccacttc	acccggata	29105	
cctcccgctc	cccttggatc	gctgctaca	actgcagttt	ggtgttcaga	cacgggttct	29165
taggtgacag	tatccacacg	gcctttgc	cggccaggcg	ctggtcgtta	atgctcacaa	29225
atccctccga	cacgtcctcc	agacacacgg	tggaatccaa	ggcgcccgtc	tacaaaacaa	29285
acacagtcat	gcttccacg	ggttctctcc	tcggtcgtac	tgcgccagcg	tgaacggcg	29345
atggtgctcc	atcagggctc	gcagcaaccg	ctgtcgccgc	ggctccggcg	29405	
aaaagcgccc	cgtctggag	tgctattcaa	aaaacgcacc	gccttatca	acagtctct	29465
cgtgcggcgg	gcccggcgg	gcacctggat	ctctgtcagg	tcttacaat	aggtacagcc	29525
catcaccacc	atgttggta	aaatccaaa	gctaaacacg	ctccacccaa	atgacatgaa	29585
ttccagcacc	gccgcggcgt	ggccatcata	caatatgcgg	aggtaaatca	ggtgccgccc	29645
cctaatacaa	acgctccca	tatacatcac	ctccttaggc	agttgataat	taaccacctc	29705
ccggtaccag	ggaaacctca	cgtttactaa	agccccaaac	accaacattt	taaaccagtt	29765
agccagcacc	accctcccg	ccttacactg	cagcgacccc	ggctgtttac	aatgacagtg	29825
aatcaccac	ctctcatacc	ccctaattgac	ctggcgtggc	tccacatcta	tagtagcaca	29885
gcacacgcac	accctcatgt	aatgcttcat	cacaaatctt	tcccaagggg	ttagtatcat	29945

gtcccagggt acgggccact cctgcagcac ggtgaaaggt acgcaggcgg gaacagtcc 30005
cacctcgac acataatgca tattcagatg ttcacactct aaaaccccg ggcttccctc 30065
caacgcagcc actggcaagt tctcagaggg tgggttaagg cggtggtgct gatagggact 30125
caatctgtgt cgacaccgtc tgtegcgttg catcgtagac caacgcttgg cgcaccgcct 30185
cgtacttcgc ccaaagaaaa cgggtgcgac gccaaacacac ttccgcgtac cgtgggttcc 30245
gcactcgagc tcgctcagtt ctcaacgcac aatgcagcca ttccctgtaat ccacacaaca 30305
gtcgcteggc ttccaaagag atgtgcacct cgtatcttat aacgtcccga tatatatcca 30365
agcaggcagt cagggccact tgcaaccagt gcacgcaggc ggactgatcg cgacacactg 30425
gaggtggagg gagagacgga agaggcatgt tactccagac ggtcgaaaag cggatcaaag 30485
tgcagatcgc gaagatggca gcgatccccg ccgctacgct ggtgatagat cacagccagg 30545
tcaaacataa tgcggtttc caaatgacct attaccgcct ccaccagagc cgccacgcgc 30605
acttccagaa acaccagcac ggctacggca ttctcctcaa aatcttcaaa cattaagctg 30665
catgattgaa tcaccccaa ataattctcc tccttccatt ctgcaaaaat ttgagtaaaa 30725
accttcgcga gattagctcc gtggcggtca aaaaggtcac ttagagcgcc ctccaccgc 30785
atgcgcaagc acaccctcat gattaaaaa tgccagtctc ctgaaccacc tgcaagttgat 30845
ttaaaagacc tatatttagga tcaattccac tctcccgag ctccacgcgt agcattagct 30905
gcaaaaagtc atttaaatct tcgcaaacta ggcggtaag ctgcggccg ggaattaggt 30965
ctgaagcagt caccacacac ataatttcca gtgaaggagt cagtctaagc agcaaaaagc 31025
cgcatgagca gtgttgaaaa ggaggggtca cgcaatgtaa catatgcagc caaaaatctc 31085
caaggtgtct gtgcataaac tccaccactg aaaagtccaa atcatgtaaa tatgccatca 31145
ccgcctcagg aaccaccacg gacacaaaaa cggccgtag caaatacatg gtgtcctgca 31205
aagcaaaaac acatttatac catagaggcg cgaattactt gggaaaaat cactcgctcc 31265
aaaactaaac aggccaccgt ctgaccgcgc cagccataaa aaaagcggtt cgaatgatta 31325
aaaagaataa tagacacaccc tcaccaggta ctggctgca actcgtgcgc ccctatcaaa 31385
accccgccga cgttcatgtc ggccatagaa aaaatgcggc ccaaataatcc caccggaatc 31445
tcacggcca gctgcagtga tagaaaaaaga acgccatgag gagcaatcac aaaatttca 31505
ggcgataaaa gcacataaaag gttagaatag ccctgctgca caggtataaa agcccgcgag 31565
ctcagcaaat gcacataaac cgcttcagcc atccctgttt accgcgaaca aaaggtcac 31625

agtacacagt tactcaaccc acacgccaca cagtattat acactcctca atgccacgt 31685
caccggcccc gaacaaactc caaaagtcca aaaagtccaa aacgcccgcg taaaagcccg 31745
ccaaaacagc acttcctcat ttactctccc acagtacgtc acttccgcgc cgccgcgc 31805
cctcgccccg ccctcacccct cgcgctccac cccgcgeccc acgtcagact cccacccgc 31865
cccgccccc cgtcatccgc accccacccct cactccaccc ctaacccgc ctcctcatta 31925
tcatattggc accgttccca aataaggtat attatgatga tg 31967

<210> 13
<211> 508
<212> PRT
<213> Simian adenovirus

<400> 13

Met Arg Arg Ala Val Gly Val Pro Pro Val Met Ala Tyr Ala Glu Gly
1 5 10 15

Pro Pro Pro Ser Tyr Glu Thr Val Met Gly Ala Ala Asp Ser Pro Ala
20 25 30

Thr Leu Glu Ala Leu Tyr Val Pro Pro Arg Tyr Leu Gly Pro Thr Glu
35 40 45

Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Tyr Asp Thr
50 55 60

Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Ile Ala Ser Leu
65 70 75 80

Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Val Gln Asn
85 90 95

Asn Asp Phe Thr Pro Val Glu Ala Gly Thr Gln Thr Ile Asn Phe Asp
100 105 110

Glu Arg Ser Arg Trp Gly Gly Asp Leu Lys Thr Ile Leu Arg Thr Asn
115 120 125

Met Pro Asn Ile Asn Glu Phe Met Ser Thr Asn Lys Phe Arg Ala Arg
130 135 140

Leu Met Val Glu Lys Val Asn Lys Glu Thr Asn Ala Pro Arg Tyr Glu

145 150 155 160
Trp Phe Glu Phe Thr Leu Pro Glu Gly Asn Tyr Ser Glu Thr Met Thr
165 170 175

Ile Asp Leu Met Asn Asn Ala Ile Val Asp Asn Tyr Leu Glu Val Gly
180 185 190

Arg Gln Asn Gly Val Leu Glu Ser Asp Ile Gly Val Lys Phe Asp Thr
195 200 205

Arg Asn Phe Arg Leu Gly Trp Asp Pro Val Thr Lys Leu Val Met Pro
210 215 220

Gly Val Tyr Thr Asn Glu Ala Phe His Pro Asp Ile Val Leu Leu Pro
225 230 235 240

Gly Cys Gly Val Asp Phe Thr Gln Ser Arg Leu Ser Asn Leu Leu Gly
245 250 255

Ile Arg Lys Arg Met Pro Phe Gln Ala Gly Phe Gln Ile Met Tyr Glu
260 265 270

Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val Ala Lys Tyr
275 280 285

Glu Ala Ser Ile Gln Lys Ala Arg Glu Gln Gly Gln Glu Ile Arg Gly
290 295 300

Asp Asn Phe Thr Val Ile Pro Arg Asp Val Glu Ile Val Pro Val Glu
305 310 315 320

Lys Asp Ser Lys Asp Arg Ser Tyr Asn Leu Leu Pro Gly Asp Gln Thr
325 330 335

Asn Thr Ala Tyr Arg Ser Trp Phe Leu Ala Tyr Asn Tyr Gly Asp Pro
340 345 350

Glu Lys Gly Val Arg Ser Trp Thr Leu Leu Thr Thr Thr Asp Val Thr
355 360 365

Cys Gly Ser Gln Gln Val Tyr Trp Ser Leu Pro Asp Met Met Gln Asp
370 375 380

Pro Val Thr Phe Arg Pro Ser Ser Gln Val Ser Asn Tyr Pro Val Val
385 390 395 400

Gly Val Glu Leu Leu Pro Val His Ala Lys Ser Phe Tyr Asn Glu Gln
405 410 415

Ala Val Tyr Ser Gln Leu Ile Arg Gln Ser Thr Ala Leu Thr His Val
420 425 430

Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Val Arg Pro Pro Ala Pro
435 440 445

Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala Leu Thr Asp His Gly
450 455 460

Thr Leu Pro Leu Arg Ser Ser Ile Ser Gly Val Gln Arg Val Thr Ile
465 470 475 480

Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val His Lys Ala Leu Gly
485 490 495

Ile Val Ala Pro Lys Val Leu Ser Ser Arg Thr Phe
500 505

<210> 14
<211> 917
<212> PRT
<213> Simian adenovirus

<400> 14

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ala
1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
20 25 30

Arg Ala Thr Asp Thr Tyr Phe Ser Leu Gly Asn Lys Phe Arg Asn Pro
35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
50 55 60

Thr Leu Arg Phe Val Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65 70 75 80

Lys Val Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Ser
100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
115 120 125

Ala Pro Asn Pro Ser Glu Trp Lys Gly Ser Asp Asn Lys Ile Ser Val
130 135 140

Arg Gly Gln Ala Pro Phe Phe Ser Thr Ser Ile Thr Lys Asp Gly Ile
145 150 155 160

Gln Val Ala Thr Asp Thr Ser Ser Gly Ala Val Tyr Ala Lys Lys Glu
165 170 175

Tyr Gln Pro Glu Pro Gln Val Gly Gln Glu Gln Trp Asn Ser Glu Ala
180 185 190

Ser Asp Ser Asp Lys Val Ala Gly Arg Ile Leu Lys Asp Thr Thr Pro
195 200 205

Met Phe Pro Cys Tyr Gly Ser Tyr Ala Lys Pro Thr Asn Glu Gln Gly
210 215 220

Gly Gln Gly Thr Asn Thr Val Asp Leu Gln Phe Phe Ala Ser Ser Ser
225 230 235 240

Ala Thr Ser Thr Pro Lys Ala Val Leu Tyr Ala Glu Asp Val Ala Ile
245 250 255

Glu Ala Pro Asp Thr His Leu Val Tyr Lys Pro Ala Val Thr Thr Thr
260 265 270

Thr Thr Ser Ser Gln Asp Leu Leu Thr Gln Gln Ala Ala Pro Asn Arg
275 280 285

Pro Asn Tyr Ile Gly Phe Arg Asp Asn Phe Ile Gly Leu Met Tyr Tyr

290

295

300

Asn Ser Thr Gly Asn Met Gly Val Leu Ala Gly Gln Ala Ser Gln Leu
305 310 315 320

Leu Met Leu Asp Ala Leu Gly Asp Arg Ser Arg Tyr Phe Ser Met Trp
340 345 350

Asn Gln Ala Val Asp Ser Tyr Asp Pro Asp Val Arg Ile Ile Glu Asn
355 360 365

His Gly Val Glu Asp Glu Leu Pro Asn Tyr Cys Phe Pro Leu Gly Gly
370 375 380

Ser Leu Val Thr Glu Thr Tyr Thr Gly Leu Ser Pro Gln Asn Gly Ser
 385 390 395 400

Asn Thr Trp Thr Thr Asp Ser Thr Thr Tyr Ala Thr Arg Gly Val Glu
405 410 415

Ile Gly Ser Gly Asn Met Phe Ala Met Glu Ile Asn Leu Ala Ala Asn
420 425 430

Leu Trp Arg Ser Phe Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp
435 440 445

Glu Tyr Lys Leu Thr Pro Asp Asn Ile Thr Leu Pro Asp Asn Lys Asn
450 455 460

Thr Tyr Asp Tyr Met Asn Gly Arg Val Ala Ala Pro Ser Ser Leu Asp
 465 470 475 480

Thr Tyr Val Asn Ile Gly Ala Arg Trp Ser Pro Asp Pro Met Asp Asn
485 490 495

Val Asn Pro Phe Asn His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser
 500 505 510

Met Leu Leu Gly Asn Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro
515 520 525

Gln Lys Phe Phe Ala Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr
530 535 540

Thr Tyr Glu Trp Asn Phe Arg Lys Asp Val Asn Met Ile Leu Gln Ser
545 550 555 560

Ser Leu Gly Asn Asp Leu Arg Val Asp Gly Ala Ser Val Arg Phe Asp
565 570 575

Ser Ile Asn Leu Tyr Ala Asn Phe Phe Pro Met Ala His Asn Thr Ala
580 585 590

Ser Thr Leu Glu Ala Met Leu Arg Asn Asp Thr Asn Asp Gln Ser Phe
595 600 605

Asn Asp Tyr Leu Cys Ala Ala Asn Met Leu Tyr Pro Ile Pro Ala Asn
610 615 620

Ala Thr Ser Val Pro Ile Ser Ile Pro Ser Arg Asn Trp Ala Ala Phe
625 630 635 640

Arg Gly Trp Ser Phe Thr Arg Leu Lys Thr Lys Glu Thr Pro Ser Leu
645 650 655

Gly Ser Gly Phe Asp Pro Tyr Phe Thr Tyr Ser Gly Ser Ile Pro Tyr
660 665 670

Leu Asp Gly Thr Phe Tyr Leu Asn His Thr Phe Lys Lys Val Ser Ile
675 680 685

Met Phe Asp Ser Ser Val Ser Trp Pro Gly Asn Asp Arg Leu Leu Thr
690 695 700

Pro Asn Glu Phe Glu Ile Lys Arg Thr Val Asp Gly Glu Gly Tyr Asn
705 710 715 720

Val Ala Gln Cys Asn Met Thr Lys Asp Trp Phe Leu Ile Gln Met Leu
725 730 735

Ser His Tyr Asn Ile Gly Tyr Gln Gly Phe Tyr Val Pro Glu Gly Tyr
740 745 750

Lys Asp Arg Met Tyr Ser Phe Phe Arg Asn Phe Gln Pro Met Ser Arg
755 760 765

Gln Val Val Asp Thr Thr Tyr Thr Asp Tyr Lys Asn Val Thr Leu
770 775 780

Pro Phe Gln His Asn Asn Ser Gly Phe Val Gly Tyr Met Gly Pro Thr
785 790 795 800

Met Arg Glu Gly Gln Ala Tyr Pro Ala Asn Tyr Pro Tyr Pro Leu Ile
805 810 815

Gly Lys Thr Ala Val Pro Ser Leu Thr Gln Lys Lys Phe Leu Cys Asp
820 825 830

Arg Thr Met Trp Arg Ile Pro Phe Ser Ser Asn Phe Met Ser Met Gly
835 840 845

Ala Leu Thr Asp Leu Gly Gln Asn Met Leu Tyr Ala Asn Ser Ala His
850 855 860

Ala Leu Asp Met Thr Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu
865 870 875 880

Leu Tyr Val Leu Phe Glu Val Phe Asp Val Val Arg Ile His Gln Pro
885 890 895

His Arg Gly Val Ile Glu Ala Val Tyr Leu Arg Thr Pro Phe Ser Ala
900 905 910

Gly Asn Ala Thr Thr
915

<210> 15
<211> 607
<212> PRT
<213> Simian adenovirus

<400> 15

Met Lys Arg Ala Arg Leu Asp Asp Asp Phe Asn Pro Val Tyr Pro Tyr
1 5 10 15

Asp Thr Pro Asn Ala Pro Ser Val Pro Phe Ile Thr Pro Pro Phe Val

20

25

30

Ser Ser Asp Gly Leu Gln Glu Lys Pro Pro Gly Met Leu Ser Leu Asn
 35 40 45

Tyr Gln Asp Pro Ile Thr Thr Gln Asn Gly Ala Leu Thr Leu Lys Leu
 50 55 60

Gly Ser Gly Leu Asn Ile Asn Gln Asp Gly Glu Leu Thr Ser Asp Ala
 65 70 75 80

Ser Val Leu Val Thr Pro Pro Ile Thr Lys Ala Asn Asn Thr Ile Gly
 85 90 95

Leu Ala Phe Asn Ala Pro Leu Thr Leu Gln Ser Asp Thr Leu Asn Leu
 100 105 110

Ala Cys Asn Ala Pro Leu Thr Val Gln Asp Asn Arg Leu Gly Ile Thr
 115 120 125

Tyr Asn Ser Pro Leu Thr Leu Gln Asn Ser Glu Leu Ala Leu Ala Val
 130 135 140

Thr Pro Pro Leu Asp Thr Ala Asn Asn Thr Leu Ala Leu Lys Thr Ala
 145 150 155 160

Arg Pro Ile Ile Thr Asn Ser Asn Asn Glu Leu Thr Leu Ser Ala Asp
 165 170 175

Ala Pro Leu Asn Thr Ser Thr Gly Thr Leu Arg Leu Gln Ser Ala Ala
 180 185 190

Pro Leu Gly Leu Val Asp Gln Thr Leu Arg Val Leu Phe Ser Asn Pro
 195 200 205

Leu Tyr Leu Gln Asn Asn Phe Leu Ser Leu Ala Ile Glu Arg Pro Leu
 210 215 220

Ala Leu Thr Thr Gly Ser Met Ala Met Gln Ile Ser Gln Pro Leu
 225 230 235 240

Lys Val Glu Asp Gly Ser Leu Ser Leu Ser Ile Glu Ser Pro Leu Asn
 245 250 255

Leu Lys Asn Gly Asn Leu Thr Leu Gly Thr Gln Ser Pro Leu Thr Val
260 265 270

Thr Gly Asn Asn Leu Ser Leu Thr Thr Thr Ala Pro Leu Thr Val Gln
275 280 285

Asn Asn Ala Leu Ala Leu Ser Val Leu Leu Pro Leu Arg Leu Phe Asn
290 295 300

Asn Thr Ser Leu Gly Val Ala Phe Asn Pro Pro Ile Ser Ser Ala Asn
305 310 315 320

Asn Gly Leu Ser Leu Asp Ile Gly Asn Gly Leu Thr Leu Gln Tyr Asn
325 330 335

Arg Leu Val Val Asn Ile Gly Gly Leu Gln Phe Asn Asn Gly Ala
340 345 350

Ile Thr Ala Ser Ile Asn Ala Ala Leu Pro Leu Gln Tyr Ser Asn Asn
355 360 365

Gln Leu Ser Leu Asn Ile Gly Gly Leu Arg Tyr Asn Gly Thr Tyr
370 375 380

Lys Asn Leu Ala Val Lys Thr Asp Ser Phe Arg Gly Leu Glu Ile Asp
385 390 395 400

Ser Asn Gln Phe Leu Val Pro Arg Leu Gly Ser Gly Leu Lys Phe Asp
405 410 415

Gln Tyr Gly Tyr Ile Ser Val Ile Pro Pro Thr Val Thr Pro Thr Thr
420 425 430

Leu Trp Thr Thr Ala Asp Pro Ser Pro Asn Ala Thr Phe Tyr Asp Ser
435 440 445

Leu Asp Ala Lys Val Trp Leu Ala Leu Val Lys Cys Asn Gly Met Val
450 455 460

Asn Gly Thr Ile Ala Ile Lys Ala Leu Lys Gly Thr Leu Leu Gln Pro
465 470 475 480

Thr Ala Ser Phe Ile Ser Phe Val Met Tyr Phe Tyr Ser Asn Gly Thr
485 490 495

Arg Arg Thr Asn Tyr Pro Thr Phe Glu Asn Glu Gly Ile Leu Ala Ser
500 505 510

Ser Ala Thr Trp Gly Tyr Arg Gln Gly Asn Ser Ala Asn Thr Asn Val
515 520 525

Thr Ser Ala Val Glu Phe Met Pro Ser Ser Thr Arg Tyr Pro Val Asn
530 535 540

Lys Gly Thr Glu Val Gln Asn Met Glu Leu Thr Tyr Thr Phe Leu Gln
545 550 555 560

Gly Asp Pro Thr Met Ala Ile Ser Phe Gln Ala Ile Tyr Asn His Ala
565 570 575

Leu Glu Gly Tyr Ser Leu Lys Phe Thr Trp Arg Val Arg Asn Arg Glu
580 585 590

Arg Phe Asp Ile Pro Cys Cys Ser Phe Ser Tyr Ile Thr Glu Glu
595 600 605

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/016614

International filing date: 15 June 2004 (15.06.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/575,429
Filing date: 28 May 2004 (28.05.2004)

Date of receipt at the International Bureau: 06 October 2004 (06.10.2004)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse